BN
Sur
1

ウィルクス

数理統計学

〈増訂新版〉

1

田中英之
岩本誠一 訳

東京図書株式会社

はしがき

本書は，大学初等程度の数学知識を持つ読者を対象に，数理統計学を導入する目的で書かれている．読者は確率および統計の予備知識を必要とはしないが，少しでも学んでいればそれにこしたことはない．本書は，第 2 次世界大戦中，プリンストン大学で講義した原稿をもとにして，さらに推敲し再編成してある．その一部の初版は 1943 年プリンストン大学（Princeton University Press）から「数理統計学」(Mathematical Statistics) として出版されている[*]．

数理統計学およびその応用の分野は，この 25 年間にめざましい速度で発展してきた．その発展の根源はおびただしいばかりの新概念の導入にある．この時代の研究成果は数種の専門誌に発表されてきたが，重要な論文が各分野の専門誌にいまも続々と発表されている．

これらの豊富な文献の主な研究成果を詳細に取り上げるべきであろうが，本書ではあえてそれを行なわず，筆者自身の先入感と好みで，数理統計学の基本的な事項と思われるものを選んで取り上げた．すなわち，本書では，近代確率論の枠の中で，細かすぎない程度に，数理統計学上の重要な結果とともに，古典的な結果を統一的にあるいは体系的に示すように，話題を選択してある．それゆえに，あるトピックスはその専門家にとっては不十分であったり，省略してあると思われるであろう．しかし，ほんの少ししかふれなかったり，詳細な解説のないトピックスに興味を持たれた読者がさらに深く研究できるように，多くの文献を引用してある．数理統計学の予備知識がなくても，本書の内容およびその一部分を詳しく体系的に学びたい読者で，数学的な理解力のある人は，本文中および巻末に掲げた文献によって，さらに研究を進めることができるであろう．

また，全体として，400 題以上の練習問題が各章の終わりに収録されている．これは，数理統計学の内容の理解を深める意図である．練習問題は追加すべきトピックスの導入とスペースの都合で本文中で十分に論ずることができなかった重要な理論を手短にまとめた

---

[*] 邦訳　小河原正己訳『数理統計学』春日出版社，1951．(訳注)

ものが大多数である.

読者の中には，本文中の数学的結果とその結果を基礎にした統計学的方法論との関係をなぜ詳細に論じないのかという疑問を持たれる方もいるであろう．本書では，その部分は最小限にとどめるようにした．統計学的方法論とその応用に関する考察は，その基本をなす数学的理論に劣らず重要である．しかし，経験上，統計学の 2 つの面すなわち数学的理論とその理論を用いた統計学的方法論は，研究論文や特定のトピックスを扱った本の中では，非常に効果的に結び合わされているように思われる．したがって，本書のように包括的に統計学を扱う場合に，両面のトピックスに同じ力を入れる愚はとるべきでないと思う．本書で試みているように，十分に議論をし，例題で基本概念を理解しながら，数理統計学上の基礎および統計学上の広範囲にわたるトピックスの基礎となる数学的理論を紹介する方法が一番適当であると思う．このアプローチは，統計学的方法論を十分理解するのになくてはならない方法であると信ずる．いうまでもなく，筆者の一番満足のいく方法である.

近代数理統計学はそのほとんどが確率論に依存している．それゆえ，実際に基礎理論を構築せずに，近代確率論の基礎理論の上で数理統計学を展開する方法がとられている．しかし，これは本書の範囲外であり，またその必要はないだろう．なぜなら，読者が確率論をさらに学ぼうとする場合，本書のどこかで引用されている事項に関して，秀れた本がすでにいくつか出版されているからである.

本書のように広範なトピックスを取り上げている書物では，全体を通じてその記号や用語を統一することは容易でない．記号や用語のいくつかは数理統計学では親しみがないように見えるが，これは他の個所との矛盾をなくしたり，あるいは，数理統計学の文献ではこれまでに体系的に取り上げられたことのないトピックスに関してまったく新しい記号や用語を導入するときに，矛盾を起こさないためである.

順序統計およびノンパラメトリック統計に関する部分はフルブライト研究生としてケンブリッジ大学に在籍していた 1951 年春から夏にかけて書いたものである．その時期に筆者は順序統計およびノンパラメトリック統計に関する書物を書こうとしていた．仕事として，書く準備と調査の機会を得たことは非常に恵まれていた．しかし，これらを数理統計学の包括的な書物にまとめようと決めたのは，かなりたってからである．この情熱を傾けた仕事は，Office of Naval Research から受けた援助なしにはなし得なかったであろう．この場をかりて深く感謝の意を表したい.

本書を書き上げるまでに研究所時代およびプリンストン在学中の友人から有益な助言と激励を受けた．特に J. W. Tukey 氏からは多岐にわたる忠告と助言をいただいたし，F. J. Anscombe 氏との論議も非常に役に立った．V. S. Varadarajan 氏には確率論の基本的な点について御指摘を受けた．D. M. Brown, D. A. Freedman, I. Cuttman, A. T. James, F. M. Sand の各氏には原稿に目を通していただき，有益な御指摘を受けた．また，D. R. Brillinger 氏には 350 題もの練習問題を作成していただいた．J. A. Hartigan 氏には証明に目を通していただき，最終校正まで手伝っていただいた．これら各氏に心から感謝の意を表したい．同僚および友人から非常に多くの助言を得たが，そのすべてを包含することはできなかった．本書の誤りや不備な点は筆者個人の責任であり，その点については読者の忌憚のない御指摘を受けたい．最後に，タイプと校正を担当していただいた Rebecca Werkman, Emily Sorenson 両女史に謝辞を贈る.

1961 年 11 月　　サムエル・S・ウィルクス

第 2 版の出版にあたって

読者から第 1 版の誤りおよび不備な点の御指摘があったので，第 2 版ではそれを改訂した．特に D. R. Cox, P. C. Fishburn, I. Guttman, E. J. Hannan, W. Hoeffding, S. Kullback, M. Kupperman, G. P. Patil の各氏から送られた誤りの御指摘に感謝したい.

1963 年 4 月　　サムエル・S・ウィルクス

目　次

第2巻の目次

第1章　序　　論

1.1　標本空間と事象

数理統計学は確率論の上に成り立っている．すなわち，数理統計学の表現と論理に厳密性を持たせるためには，測度論と積分論に負うところが大きいのである．本書では，確率の測度論的な表現に Kolmogorov (1933 *a*) の成果を用いる．数理統計学に確率論を用いるときには，使いやすい形で，確率論の基本的な定義，概念，しくみを示す必要がある．確率論の詳細および広範な部分に興味のある読者は Doob (1953), Gnedenko と Kolmogorov (1954), Feller (1957), Kolmogorov (1933 *a*), Lévy (1925, 1937), Loève (1955) の本を参照すると良い． 同じく測度論に関しては，Halmos (1950), Munroe (1953) を参照されたい．

本書で必要な確率論の基礎として，標本空間と事象の概念，事象の表現および事象の結合や分割をまず取り上げる．なお，これらの概念を取り扱うのに便利な集合論を用いる．

要素 e の集合を R で表わし，**標本点**，**事象点**，または単に**点**と呼ぶ．標本点の数は有限でも無限でも良い． R は**標本空間**または**出現空間**とも呼ばれるが，本書では前者を使う．標本点は与えられた条件の集合の下での試行，実験，操作の起りうる結果と考えて良い．しかし，ここでは前もって定義を与えないことにしよう．標本空間 R は単に与えられた条件下で試行がなされたとき見られるすべての起りうる結果の集合である．

例題　コインを 1 回投げたとき， 上を向いている面を e で表わすと， 標本空間 R は表，裏の 2 つの標本点から成り立つ．コインを 2 回投げると，e は上を向いた面の組合せを表わし，標本空間 R は 4 つの標本点，HH, HT, TH, TT からなる．そして，実際に 2 回投げると，これらの標本点のうちどれか 1 つが出現する．
13 枚のトランプを配り，任意の持札を e で表わせば，標本空間 R は $\binom{52}{13}$ 個の標本点

からなり，各標本点は持札の 1 つとなっている．そして，13 枚のカードの持札を実際に配ると $\binom{52}{13}$ のうちどれか 1 つが出現する．

電球が“切れる”までついていたとする．e を電球がともっていた時間の長さとすれば，e は正の数で，標本空間 R はすべての正の実数からなっている．k 個の電球 $B_1, B_2, \cdots, B_k$ のすべてが“切れる”まで試行を続けると，e は，電球 $B_1, B_2, \cdots, B_k$ の寿命を表わす k 個の正の数 $(x_1, x_2, \cdots, x_k)$ の集合となり，標本空間 R は，座標系がすべて正の k 次元ユークリッド空間の点からなる．

これらの例題で試行がなされ，その結果が標本空間 R に属する 1 つの標本点 e として表現できる．

標本点 e は根元事象ともいい，標本点の集合 E は R の点の部分集合で，事象という．事象はもちろん 1 点のみでできていることもある．事象 E が起るとは，標本点 e（実験，試行，操作の結果を表わす）が E に含まれていることを意味している．一般には，このような個々の標本点よりも事象と事象の族の方が重要である．また，個々の標本点に関する確率よりも事象に関する確率の方が重要である．そして最も重要な事象は可測事象すなわち確率が定義されている集合である．可測の概念は 1.4 節で論ずる．

集合論では，R の部分集合のある族と，このような族に定義された確率（集合関数）が重要である．しかしながら標本空間 R の点集合は事象といわれる．実際には“集合”と“事象”を同じ用語として使うと便利である．

この理論を進める前に集合代数の基本的概念を導入しよう．

1.2　事象の結合，分割に関する定義と法則

本節では，集合 $E, E_1, E_2, \cdots$ は標本空間 R 内の標本点よりなる事象を表わす．このような集合は e 集合ともいわれる．n 個の集合の有限列は $E_1, \cdots, E_n$ と書き，可付番無限列は $E_1, E_2, \cdots$ と表わす．

e が E の要素ならば次のように表わす．

(1.2.1)
$$e \in E.$$

そして e は E に属するという．

E_1 と E_2 が集合で，E_1 のどの点も E_2 に含まれる場合，E_1 は E_2 の部分集合といい，次のように表わす．

(1.2.2)
$$E_1 \subset E_2 \qquad \text{または} \qquad E_2 \supset E_1.$$

もし E_1 が e という 1 点だけの場合には，$E_1 \subset E_2$ は $e \in E_2$ という表現に等しい．

2 つの集合 E_1，E_2 が等しいとは，$E_1 \subset E_2$ かつ $E_1 \supset E_2$ の場合であり，次のように表わす．

(1.2.3)
$$E_1 = E_2.$$

空集合（0 集合）は点を含まない集合のことである．この場合，次のように表わす．

(1.2.4)
$$E = \varnothing.$$

空集合 $\varnothing$ は $E \subset R$ なるあらゆる集合の部分集合となる．

E_1 と E_2 の両方に含まれる点集合を E_1 と E_2 の共通部分または積集合といい，次のように表わす．

(1.2.5)
$$E_1 \cap E_2,$$

そして“E_1 キャップ E_2”と読む．

(1.2.6)
$$E_1 \cap E_2 = \varnothing$$

が成り立つとき，すなわち共通点がない場合，集合 E_1 と E_2 は互いに素[*]である．

集合列 $E_1, E_2, \cdots$ の共通部分は集合が n 個の場合には

(1.2.7)
$$\bigcap_{\alpha=1}^{n} E_\alpha,$$

可付番無限の場合には

(1.2.8)
$$\bigcap_{\alpha=1}^{\infty} E_\alpha$$

で示される．さらに一般化して，任意の集合の集まり $\{E_\alpha : \alpha \in T\}$ に対して，集合の共通部分は $\bigcap_{\alpha \in T} E_\alpha$ で示され，さらに α の値域 T に対して，簡潔に $\bigcap_{\alpha} E_\alpha$ で示す．もちろん T は整数の集合とは限らず，より一般的な空間の要素の集合でも良い．本書の例題では，ほとんどの T は列（整数の有限または可付番集合）である．$E_\alpha \cap E_\beta = \varnothing$，$\alpha \neq \beta$ であれば，$E_1, \cdots, E_n$ は対ごとに素，または相互に素[**]である，という．

集合 E_1 と E_2 の少なくとも一方に含まれる点全体の集合は E_1 と E_2 の合併集合ま

---

*) 排反ともいう．（訳注）

**) 対ごとに排反，相互に排反ともいう．（訳注）

たは和集合と呼び，次のように書く*).

(1.2.9)
$$E_1 \cup E_2,$$

そして "E_1 カップ E_2" と読む．E_1 が任意の e の集合であって，かつ $E_2=\varnothing$ ならば $E_1 \cup E_2 = E_1$ となる．

有限集合列 $E_1, \cdots, E_n$ の合併集合は

(1.2.10)
$$\bigcup_{\alpha=1}^{n} E_\alpha$$

で示され，n が可付番無限ならば

(1.2.11)
$$\bigcup_{\alpha=1}^{\infty} E_\alpha$$

と書く．または任意の集合の集まり $\{E_\alpha : \alpha \in T\}$ に対しては $\bigcup_{\alpha} E_\alpha$ と書く．

E_2 に含まれない E_1 の点のすべてからなる集合は E_1 と E_2 の差といい，

(1.2.12)
$$E_1 - E_2$$

と書く．

$E_1 \subset E_2$ ならば $E_1 - E_2$ は空集合であり，また $E_1 \cap E_2 = \varnothing$ ならば $E_1 - E_2 = E_1$ である．$E_2 \subset E_1$ のとき，$E_1 - E_2$ は E_1 と E_2 の**真の差**と呼ばれ，この場合 $E_1 - E_2$ は E_1 から E_2 を減じて得られる．

真の差 $E_1 - E_2$ は，時には E_1 に関する E_2 の**補集合**ともいわれ，次のように書く．

(1.2.13)
$$E_1 - E_2 = \overline{E_2}(E_1)$$

特殊な場合ではあるが，E_1 が標本空間 R の全体であるときは，$R - E_2$ は E_2 の補集合と呼び，次のように書く．

(1.2.14)
$$R - E_2 = \overline{E}_2.$$

一般的には

(1.2.15)
$$E_1 - E_2 = E_1 \cap \overline{E}_2.$$

集合の合併集合，共通部分，差の概念を図で説明するのにベン図を用いる．標本空間は図 1.1 の長方形の内部の点集合で表わされるものとし，E_1 と E_2 はそれぞれ大きい円と小さい円の内部の点集合とする．合併集合 $E_1 \cup E_2$ は少なくとも 2 つの円のどちら

*) 合併集合（和集合）$E_1 \cup E_2$ は，時には $E_1 + E_2$ と書かれ，共通部分（積集合）$E_1 \cap E_2$ は $E_1 \cdot E_2$ と書かれることもある．

か一方に含まれている点集合で，共通部分 $E_1 \cap E_2$ は両方の円の内部にある点集合である．差 $E_1 - E_2$ $(= E_1 \cap \overline{E_2})$ は大きい円の内部にあり，小さい円の内部にはない点集合 $E_2 - E_1$ $(= \overline{E_1} \cap E_2)$ も同様に解釈できる．$\overline{E_1} \cap \overline{E_2} = R - (E_1 \cup E_2)$ はどちらの円にも含まれない矩形内部の点集合である．

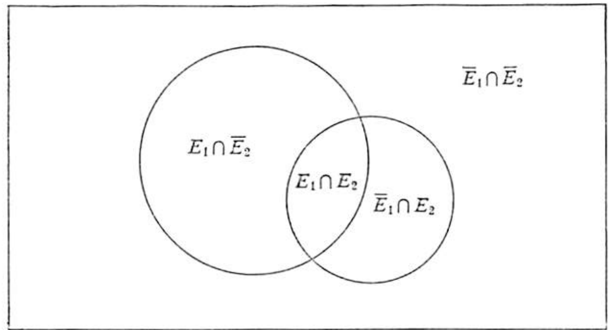

図 1.1　E_1 と E_2 によって生成される 4 つの基本的な互いに素な集合 $E_1 \cap E_2$, $E_1 \cap \overline{E_2}$, $\overline{E_1} \cap E_2$, $\overline{E_1} \cap \overline{E_2}$ を示すベン図．

これらの集合論の概念を事象に合わせて説明すると，次のような解釈ができる*).

(i) $E_1 \subset E_2$ は E_1 が起れば E_2 が起ることを意味している．すなわち標本点 e は E_2 で起らずに E_1 で起ることはありえないということである．

(ii) $E = \varnothing$（空集合）は事象 E は起りえないことを意味する．

(iii) $E = R$（標本空間全体）は事象 E は必ず起ることを意味する．

(iv) 共通部分 $E_1 \cap E_2$ は事象 E_1 と E_2 が共に起るような事象である．

(v) 合併集合 $E_1 \cup E_2$ は事象 E_1 と E_2 の少なくとも一方が起るような事象である．

(vi) 差 $E_1 - E_2$ は E_1 が起り E_2 が起らないような事象である．

合併集合や積集合をつくる演算が可換であることは容易に確かめられる．すなわち

(1.2.16)
$$E_1 \cup E_2 = E_2 \cup E_1, \qquad E_1 \cap E_2 = E_2 \cap E_1,$$

さらに結合的でもある．すなわち

*) (1.2.1) から (1.2.15) までの定義や演算はすべて標本点（事象）の集合だけに用いられるものではなく，あらゆる種類の要素やその集合にも用いられることに注目すべきである．その要素は点集合であるかも知れないし，人々や自動車等でもよい．たとえば $\mathscr{F}$ が集合 E を含む集合の集まりだとすれば $E \in \mathscr{F}$ と書くし，その集まり $\mathscr{F}$ の中のどの集合も集まり $\mathscr{G}$ の中に含まれるとすれば，$\mathscr{F} \subset \mathscr{G}$ のように書くといった具合である．

(1.2.17)　$(E_1 \cup E_2) \cup E_3 = E_1 \cup (E_2 \cup E_3)$, $(E_1 \cap E_2) \cap E_3 = E_1 \cap (E_2 \cap E_3)$,
分配的でもある．すなわち

(1.2.18)　$$E_1 \cap (E_2 \cup E_3) = (E_1 \cap E_2) \cup (E_1 \cap E_3).$$

E_1 と真の差 $E_2 - (E_1 \cap E_2)$ が互いに素であることもまた明らかであり，

(1.2.19)　$$E_1 \cup E_2 = E_1 \cup [E_2 - (E_1 \cap E_2)].$$

同様に，$E_1 \cap E_2$ と真の差 $E_1 - (E_1 \cap E_2)$，$E_2 - (E_1 \cap E_2)$ の 3 つは互いに素である．したがって

$$E_1 \cup E_2 = (E_1 \cap E_2) \cup [E_1 - (E_1 \cap E_2)] \cup [E_2 - (E_1 \cap E_2)].$$

R の任意の 2 つの部分集合 E_1, E_2 の共通部分は，合併集合と集合の差を用いて次のような形で表わせる．

(1.2.20)　$$E_1 \cap E_2 = R - (\overline{E_1} \cup \overline{E_2}).$$

同様に合併集合については下のようになる．

(1.2.21)　$$E_1 \cup E_2 = R - (\overline{E_1} \cap \overline{E_2}).$$

また次のような関係も成り立つ．

(1.2.22)　$$(E_1 \cap E_2) \subset E_\alpha, \qquad \alpha = 1, 2,$$

(1.2.23)　$$E_\alpha \subset (E_1 \cup E_2), \qquad \alpha = 1, 2.$$

前述の結果のいくつかを有限または無限の集合の集まりに拡張しておくと，後節で役に立つ．証明は簡単なので読者に残しておく．

1.2.1　任意の集合の集まり $\{E_\alpha : \alpha \in T\}$ に対して，

(1.2.20 a)　$$\bigcap_\alpha E_\alpha = R - \bigcup_\alpha \overline{E_\alpha}$$

(1.2.21 a)　$$\bigcup_\alpha E_\alpha = R - \bigcap_\alpha \overline{E_\alpha}.$$

(1.2.19) の拡張は次のようになる．

1.2.2　$E_1, E_2, \cdots$ を集合列（有限でも無限でもよい）とすれば，E_1 と真の差 $E_2 - (E_2 \cap E_1), E_3 - [E_3 \cap (E_2 \cup E_1)], \cdots, E_n - [E_n \cap (E_{n-1} \cup \cdots \cup E_1)], \cdots$ は互いに素な集合であり，それらの合併集合は $\bigcup_\alpha E_\alpha$ である．

1.2.2 は集合列の合併集合を，互いに素な集合列に分割する簡単な法則を与えている．E_1 と差 $E_2 - E_1, E_3 - E_1 \cup E_2, \cdots, E_n - E_1 \cup \cdots \cup E_{n-1}, \cdots$ もまた互いに素な集合で，

それらの合併集合は $\bigcup_\alpha E_\alpha$ となることも注意しなければならない．

$E_1, E_2, \cdots$ が R における可付番無限集合列ならば，この列の中の無限に多くの集合に属する点全体からなる集合 E^* は，その列の**上極限**と呼び，次のように書く．

(1.2.24)　$$E^* = \limsup_\alpha E_\alpha.$$

同様に，その列の中の有限個をのぞいた集合のすべてに属する点全体からなる集合 E_* は，その列の**下極限**と呼び，次のように書く．

(1.2.25)　$$E_* = \liminf_\alpha E_\alpha.$$

$E^* = E_*$ ならば，列 $E_1, E_2, \cdots$ の**極限**が存在するといい，次のように表わす．

(1.2.26)　$$\lim_{\alpha\to\infty} E_\alpha \qquad \text{または} \qquad \lim_\alpha E_\alpha.$$

拡大列または**増加列** $E_1 \subset E_2 \subset \cdots$ に対して，列の極限が存在して，すべての集合 $\{E_\alpha\}$ の合併集合に等しい．すなわち，

(1.2.27)　$$\lim_\alpha E_\alpha = \bigcup_\alpha E_\alpha.$$

縮少列または**減少列** $E_1 \supset E_2 \supset \cdots$ に対しては，極限が存在して，すべての集合の共通部分に等しい．よって，

(1.2.28)　$$\lim_\alpha E_\alpha = \bigcap_\alpha E_\alpha.$$

R の中の任意の集合列 $E_1, E_2, \cdots$ の場合には

$$\limsup E_\alpha = \bigcap_{n=1}^{\infty} \bigcup_{\alpha=n}^{\infty} E_\alpha, \qquad \liminf E_\alpha = \bigcup_{n=1}^{\infty} \bigcap_{\alpha=n}^{\infty} E_\alpha.$$

拡大列と縮少列は確率論と統計学で最も重要な列である．$E_1, E_2, \cdots$ が拡大または縮少集合列のいずれかである場合には，便宜上，**単調列**と呼ぶ．単調列の極限は常に存在する．その極限を E で表わせば，列が拡大か縮少かによって，それぞれ $E = \bigcup_\alpha E_\alpha$ または $E = \bigcap_\alpha E_\alpha$ となる．後の節や章を見れば，この単調列が確率論で基礎的な役割を果たしていることがわかるだろう．

1.3　集　合　体

前述のように，標本空間 R の個々の標本点よりもさらに重要なことは，R の中の部分集合（事象）の族である．特に，**集合体**すなわち下に示す規則を満足する集合の族であ

る．

R の空でない集合族 $\mathscr{F}$ が次の性質を満足するとき，**ブール集合体**と呼ぶ．

A1　　$E \in \mathscr{F}$ ならば $\overline{E} \in \mathscr{F}$.

A2　　$E_1 \in \mathscr{F}$, $E_2 \in \mathscr{F}$ ならば $E_1 \cup E_2 \in \mathscr{F}$.

上の性質に加えて次の性質も満足するとき，集合族 $\mathscr{F}$ は**ボレル集合体**[*)]と呼ばれる．

A3　　$E_1, E_2, \cdots$ が $\mathscr{F}$ に属する可付番無限集合列ならば，$\bigcup_{\alpha} E_\alpha \in \mathscr{F}$.

実際には，**A2** はボレル集合体には不必要である．なぜならば **A2** は **A3** に含まれるからである．**A2** の中で $E_2 = \overline{E_1}$ となるものをとれば，$E_1 \cup \overline{E_1} = R$ となり，$R \in \mathscr{F}$ となることがわかる．また A_1 において E を R にすると，$\overline{E}$ は空集合となるから，空集合はブール集合体にもボレル集合体にも常に含まれていることになる．

A2 を逐次用いると，ブール集合体 $\mathscr{F}$ に属するあらゆる有限個の集合の合併集合もまた $\mathscr{F}$ に属することがわかる．**A1** と **A2** および (1.2.20) から，ボレル集合体の中の有限または可付番無限集合列の共通部分もまた $\mathscr{F}$ に属することも証明できる．

もし R が N 個の（異なる）有限個の標本点を含んでいれば，すべての可能な事象の族は有限であって，この族はブール集合体である．実際この場合のブール集合体 $\mathscr{F}$ は次にあげる 2^N 個の集合（事象）からなっている．すなわち空集合，1 点からなる集合全体，2 点からなる集合全体，…，そして最後に R 自身（N 個の点からなる集合）．ここで，1 点からなる N 個の集合による族を $\mathscr{F}_0$ で表わせば，ブール集合体 $\mathscr{F}$ に属するどの有限個の集合も，**A1** と **A2** を有限回用いて $\mathscr{F}_0$ の集合から導くことができる．このようにブール集合体 $\mathscr{F}$ は，初期族 $\mathscr{F}_0$ から生成できる．ここでブール集合体 $\mathscr{F}$ を生成する初期族 $\mathscr{F}_0$ を選ぶには，多くの異なった方法があることを強調しておこう．すでに述べた 1 点からなる集合の族はおそらくその最も簡単なものであろう．

有限な集合族から生成されるブール集合体は，もちろん，有限個の事象しか含んでいない．しかし R が無限個の標本点 —— というよりむしろ可付番無限個の —— を含んでいる場合には，ブール集合体は多くの重要な問題を取り扱うには不適当な事象族といえる．

*) ブール集合体はブール代数とか有限加法的集合族とも呼ばれる．ボレル集合体は σ 代数または完全加法的集合族とも呼ばれる．

A3 は R が無限に多くの標本点を含んでいる場合に，より大きな事象族を取り扱うために導入されたものである．

さて，R は標本空間で，$\mathscr{F}_0$ は R における空でない事象族であるとする．$\mathscr{B}$ は $\mathscr{F}_0$ の中の集合をすべて含むような任意のボレル集合体とする．このようなボレル集合体の少なくとも 1 つが存在することは明らかである．なぜなら $\mathscr{B}$ を R の可能な部分集合のすべてからなる族としてもよい．もちろんそれは族 $\mathscr{F}_0$ を含んでいるからである．$\mathscr{B}(\mathscr{F}_0)$ を $\mathscr{F}_0$ を含む各ボレル集合体に属する集合全体からなる族とする．すなわち $\mathscr{B}(\mathscr{F}_0)$ は族 $\mathscr{F}_0$ を含むボレル集合体すべての共通部分とすると，$\mathscr{B}(\mathscr{F}_0)$ の中の集合は **A1**, **A2**, **A3** を満たすから $\mathscr{B}(\mathscr{F}_0)$ はボレル集合体である．

さらに，定義から，$\mathscr{B}(\mathscr{F}_0)$ は $\mathscr{F}_0$ を含むあらゆるボレル集合体に含まれている．したがって，これは $\mathscr{F}_0$ を含む唯一の最も小さいボレル集合体であり，族 $\mathscr{F}_0$ によって生成されるボレル集合体と呼ばれる．

要約すると

1.3.1　　$\mathscr{F}_0$ を R における空でない集合族とすると，$\mathscr{F}_0$ を含む任意のボレル集合体 $\mathscr{B}$ に対し $\mathscr{B}(\mathscr{F}_0) \subset \mathscr{B}$ となる唯一のボレル集合体 $\mathscr{B}(\mathscr{F}_0)$ が存在する．

R が実数軸 R_1（標本点 e が実数である）で，$\mathscr{F}_0$ を初期族としてすべての半開区間[*)] $(a, b]$ からなる族にとれば，ボレル集合体 $\mathscr{B}(\mathscr{F}_0)$ の集合は**実数直線のボレル集合**と呼ばれ，この族 $\mathscr{B}(\mathscr{F}_0)$ は $\mathscr{B}_1$ で表わされる．$\mathscr{F}_0$ を (a, b), $[a, b)$, $[a, b]$ の族，または R_1 の開集合全体か，閉集合全体の族とすれば，同様なボレル集合体が得られる．R_1 のボレル集合は，有限個または可付番無限個の区間で，有限回または可付番無限回の合併集合，共通部分，補集合，差などの操作によって得られるすべての集合を含んでいる族である．

一般的にいえば，標本空間 R が k 次元ユークリッド空間 R_k のとき，初期族 $\mathscr{F}_0$ を k 次元半開区間 $(a\,;\,b]_k$ 全体の族とすれば，生成されたボレル集合体 $\mathscr{B}(\mathscr{F}_0)$ は R_k のボ

*) 通常，$(a, b]$ を区間 $a < x \leqslant b$ と表わして用いている，(a, b), $[a, b)$, $[a, b]$ に対しても同様な表わし方が用いられる．x が k 次元実ベクトル，すなわち $x_1, \cdots, x_k$ が k 次元ユークリッド空間の 1 点である場合には，$(a_1, \cdots, a_k\,;\,b_1, \cdots, b_k]$ またはもっと簡単に $(a\,;\,b]_k$ を k 次元区間 $a_i < x_i \leqslant b_i$, $i = 1, \cdots, k$ と表わし，$(a\,;\,b)_k$, $[a\,;\,b)_k$, $[a\,;\,b]_k$ に対しても同様な表わし方をする．

レル集合と呼ばれ，$\mathscr{B}_k$ と表わされる．もちろん R_k のボレル集合を生成する初期族 $\mathscr{F}_0$ の選び方は他にもある．たとえば $(a\,;\,b]_k$，$[a\,;\,b)_k$，$[a\,;\,b]_k$ からなる族でも良い．R_k におけるボレル集合の族は，上記のどの型の区間で始めるにも，有限または可付番無限個の合併集合，共通部分，差，補集合でつくられる集合をすべて含んでいることを特に注意しておく．

1.4 確率測度

前節では事象の表現や取り扱いに関して述べた．ここではこれらの事象に関する確率を定める問題および確率を扱う法則について考えよう．事象の確率の一般概念は“与えられた一組の条件”のもとでの実験の繰り返し試行において，事象が起る相対度数の概念の抽象化である．E は基本標本空間 R における事象であり，“基本的な”実験が行なわれるたびに，その結果が R におけるある標本点 e に相当する．実験が m 回行なわれたとき，m_E をその結果の事象点 e が E に属する回数，すなわち E が起る度数とする．m 回の試行での E の相対度数は m_E/m となる．この比率は明らかに区間 $[0, 1]$ にはいっている：$E=R$ ならば必然的にそれは 1 になり，E が空集合ならば当然 0 となる．E_1 と E_2 が互いに素な事象であれば，$E_1\cup E_2$ が起る度数は $m_{E_1}+m_{E_2}$ であり，$E_1\cup E_2$ の相対度数は $(m_{E_1}+m_{E_2})/m=m_{E_1}/m+m_{E_2}/m$ である．このことは 2 つの互いに素な事象の合併集合の相対度数は，その 2 つの事象の相対度数の和に等しいことを示している．任意の有限個の互いに素な事象の相対度数と，それらの合併集合の相対度数に関しても同様に表わせる．

“基本的な”実験の試行を無限に長く続ける場合（明らかに実行不可能ではあるが！）の事象 E の相対度数を考えれば，事象の確率を導入したことと同じである．しかし，ここで論じる必要はないが，もし無限に続く試行において収束の性質がある場合に，相対度数の概念を厳密に定式化して確率論を設定しようとすると，厄介な問題が出て来る．この問題は von Mises (1931) により考えられた．その定式化は Kolmogorov (1933 a) により最初に提唱され，本質的には次のような仮定からなるものである．すなわち確率問題に関する事象を取り扱う場合に，比較的簡単な事象の初期族での事象の確率として，区間 $[0,1]$ 上の数を定める．これらの初期確率からさらに複雑な事象の確率を定める法則を用いれば，もとの確率問題に関する事象を十分に含んでいる事象族のどんな事象の確率をも定められる．実際に事象の確率を定めるには，通常，“基本的な”実験の試行を長く続けた結果得られるであろう事象の相対度数をもとにする．数学的な理論のはじまりは，コルモゴロフのこの確率論の出現以後である．もちろん，与えられたどのような状態でも確率が正しく定められるか，という疑問が出て来る．後章の統計的仮説検定の問題でこれを取り上げる．事象の族に対する確率測度，確率分布を定義すれば，これらの概念は定式化できる．

ブール集合体 $\mathscr{F}$ のすべての集合に対して定義された集合関数 P と次の 3 つの性質はブール集合体 $\mathscr{F}$ 上の確率測度を示すものである：

B1　$\mathscr{F}$ に含まれるすべての事象 E に対して，非負で実数の $P(E)$ を E に対応させ，事象 E の確率と呼ぶ．$P(E)$ は $P(e\in E)$ と書くこともある．

B2　$E_1, E_2, \cdots$ が $\mathscr{F}$ における互いに素な可付番無限集合列で，その合併集合が $\mathscr{F}$ に含まれるならば

$$P\Big(\bigcup_{\alpha=1}^{\infty}E_\alpha\Big)=\sum_{\alpha=1}^{\infty}P(E_\alpha).$$

B3　$P(R)=1.$

P がボレル集合体 $\mathscr{F}$ のすべての集合に対して定義され，かつ **B1**，**B2**，**B3** を満足する集合関数ならば，P をボレル集合体 $\mathscr{F}$ 上の確率測度と呼ぶ．この場合，ボレル集合体の定義から，もちろん $\bigcup_{\alpha=1}^{\infty}E_\alpha$ は $\mathscr{F}$ に属している．

組 $(R,\mathscr{F},P)$ を確率空間と呼ぶ．**B2** は互いに素な集合の有限個の集まり $E_1,\cdots,E_n$ に対しても成り立つ．なぜならば $E_{n+1}, E_{n+2},\cdots$ のすべてを空集合 $\varnothing$ と考えれば，$P(\varnothing)=0$ となるからである．

すべての $E\in\mathscr{F}$ に対して $P(E)$ が有限であり，かつ **B1**（非負という制限なしに）と **B2** を満足する集合関数 P は完全加法的集合関数と呼ばれる．

重要なのはあるブール集合体 $\mathscr{F}$ から生成される最小のボレル集合体 $\mathscr{B}(\mathscr{F})$ であり，この場合，$\mathscr{B}(\mathscr{F})$ は $\mathscr{F}$ を含んでいる．

ここで $\mathscr{B}(\mathscr{F})$ における事象の確率に関して，いくつかの定理を述べる．$\mathscr{B}(\mathscr{F})$ 上の確

率測度は $\mathscr{F}$ におけるすべての集合に対しても確率を与えることになるから，これらの定理はもちろん自動的に $\mathscr{F}$ に含まれる事象に対しても成り立つ．

1.4.1　E が $\mathscr{B}(\mathscr{F})$ に含まれる任意の事象であれば，$P(E)+P(\overline{E})=1$.

1.4.1 と **B3** から $P(\varnothing)=0$ となることはただちにわかる．$P(E)=0$ となるような $\varnothing$ でない事象 E は零確率の事象（または集合）と呼ばれる．

1.4.2　E_1 と E_2 が $\mathscr{B}(\mathscr{F})$ に含まれる事象で，$E_1\supset E_2$ ならば，$0\leqslant P(E_1-E_2)=P(E_1)-P(E_2)$ かつ $P(E_1)\geqslant P(E_2)$.

1.4.3　E_1 と E_2 が $\mathscr{B}(\mathscr{F})$ に含まれる事象ならば，$P(E_1\cup E_2)=P(E_1)+P(E_2-E_1\cap E_2)$. また $P(E_1\cup E_2)=P(E_1)+P(E_2)-P(E_1\cap E_2)$ でもある．

1.4.4　一般的に，$E_1,\cdots,E_n$ が $\mathscr{B}(\mathscr{F})$ に属す n 個の集合のとき

$$P\left(\bigcup_{\alpha=1}^{n}E_\alpha\right)=P(E_1)+P(E_2-E_2\cap E_1)+\cdots+P(E_n-E_n\cap[E_1\cup\cdots\cup E_{n-1}]), \tag{1.4.1}$$

また

$$P\left(\bigcup_{\alpha=1}^{n}E_\alpha\right)=\sum_{\alpha=1}^{n}P(E_\alpha)-\sum_{\beta>\alpha=1}^{n}P(E_\alpha\cap E_\beta)+\cdots+(-1)^{n-1}P(E_1\cap\cdots\cap E_n). \tag{1.4.1 a}$$

1.4.1, **1.4.2**, **1.4.3**, **1.4.4** の証明は練習問題として読者に残しておく．$\mathscr{B}(\mathscr{F})$ における無限集合列 $E_1,E_2,\cdots$ を考えると，(1.4.1) と (1.4.1 a) はそれぞれ次のようになる．

$$P\left(\bigcup_{\alpha=1}^{\infty}E_\alpha\right)=P(E_1)+P(E_2-E_2\cap E_1)+P(E_3-E_3\cap[E_1\cup E_2])+\cdots \tag{1.4.1$'$}$$

また（右辺の和が収束すると仮定して）

$$P\left(\bigcup_{\alpha=1}^{\infty}E_\alpha\right)=\sum_{\alpha=1}^{\infty}P(E_\alpha)-\sum_{\beta>\alpha=1}^{\infty}P(E_\alpha\cap E_\beta)+\sum_{\gamma>\beta>\alpha=1}^{\infty}P(E_\alpha\cap E_\beta\cap E_\gamma)-\cdots. \tag{1.4.1 $a$$'$}$$

(1.4.1) と (1.4.1 a) は，それぞれ (1.4.1)$'$ と (1.4.1 a)$'$ の特別な場合であって，

$E_{n+1},E_{n+2},\cdots$ を空集合として得られるものである．

次に述べる定理は第 2 章の基礎になる重要なものである．

1.4.5　$E_1,E_2,\cdots$ がボレル集合体 $\mathscr{B}(\mathscr{F})$ における集合の単調列ならば

$$\lim_{\alpha\to\infty}P(E_\alpha)=P\left(\lim_{\alpha\to\infty}E_\alpha\right).$$

1.4.5 の証明は，まず増加列 $E_1\subset E_2\subset\cdots$ の場合を考える．このとき $E_1,E_2-E_1,\cdots,E_\alpha-E_{\alpha-1}$, $\alpha=2,3,\cdots$ は互いに素であり，次の式が成り立つ．

$$E_\alpha=E_1\cup(E_2-E_1)\cup\cdots\cup(E_\alpha-E_{\alpha-1}). \tag{1.4.2}$$

$\alpha\longrightarrow\infty$ として極限をとれば

$$\lim_{\alpha\to\infty}E_\alpha=E_1\cup(E_2-E_1)\cup\cdots. \tag{1.4.3}$$

これに **B2** を用いて

$$P\left(\lim_{\alpha\to\infty}E_\alpha\right)=P(E_1)+P(E_2-E_1)+\cdots. \tag{1.4.4}$$

しかし

$$\begin{aligned}P(E_1)+P(E_2-E_1)+\cdots&=\lim_{\alpha\to\infty}[P(E_1)+P(E_2-E_1)+\cdots+P(E_\alpha-E_{\alpha-1})]\\&=\lim_{\alpha\to\infty}P(E_1\cup(E_2-E_1)\cup\cdots\cup(E_\alpha-E_{\alpha-1})).\end{aligned} \tag{1.4.5}$$

(1.4.2) を用いれば

$$\lim_{\alpha\to\infty}P(E_1\cup(E_2-E_1)\cup\cdots\cup(E_\alpha-E_{\alpha-1}))=\lim_{\alpha\to\infty}P(E_\alpha).$$

したがって

$$P\left(\lim_{\alpha\to\infty}E_\alpha\right)=\lim_{\alpha\to\infty}P(E_\alpha). \tag{1.4.6}$$

増加列の場合の **1.4.5** の証明はこれで終わった．

減少列 $E_1\supset E_2\supset\cdots$ は，その補集合 $\overline{E_1},\overline{E_2},\cdots$ が増加列になるから，上の証明から

$$P\left(\lim_{\alpha\to\infty}\overline{E_\alpha}\right)=\lim_{\alpha\to\infty}P(\overline{E_\alpha}), \tag{1.4.7}$$

すなわち

$$P\left(\lim_{\alpha\to\infty}(R-E_\alpha)\right)=\lim_{\alpha\to\infty}P(R-E_\alpha). \tag{1.4.8}$$

これは次のように書いてもよい．

$$P\left(R-\lim_{\alpha\to\infty}E_\alpha\right)=\lim_{\alpha\to\infty}P(R-E_\alpha). \tag{1.4.9}$$

これに **1.4.2** を用いれば (1.4.6) になる．もちろん減少列の場合は

$$\lim_{\alpha\to\infty} E_\alpha = \bigcap_{\alpha=1}^{\infty} E_\alpha.$$

次にあげる被覆定理は後節でも使用される．

1.4.6　$E, E_1, E_2, \cdots$ が $E \subset \bigcup_{\alpha=1}^{\infty} E_\alpha$ となる $\mathscr{B}(\mathscr{F})$ における事象ならば

$$P(E) \leqslant \sum_{\alpha=1}^{\infty} P(E_\alpha).$$

$E, E_1, E_2, \cdots$ が有限集合列でも無限集合列でも，$E \subset \bigcup_{\alpha=1}^{\infty} E_\alpha$ のとき，$\bigcup_{\alpha=1}^{\infty} E_\alpha$ を E の被覆という．

1.4.6 の証明は，$E \subset \bigcup_{\alpha=1}^{\infty} E_\alpha$ が次のように表わせることを利用する．

(1.4.10)　$$E = E \cap \left(\bigcup_{\alpha=1}^{\infty} E_\alpha\right).$$

1.2.2 を用いれば

(1.4.11)　$$E = [E \cap E_1] \cup [E \cap (E_2 - E_2 \cap E_1)] \cup \cdots.$$

上の式で [　] 内の集合は互いに素で，$\mathscr{B}(\mathscr{F})$ に属している．したがって

(1.4.12)　$$P(E) = P(E \cap E_1) + P(E \cap (E_2 - E_2 \cap E_1)) + \cdots.$$

しかし

$$[E \cap E_1] \subset E_1,\ [E \cap (E_2 - E_2 \cap E_1)] \subset E_2,\ \cdots$$

であるから，**1.4.2** を用いれば

(1.4.13)　$$P(E) \leqslant P(E_1) + P(E_2) + \cdots.$$

1.4.6 の証明はこれで終わる．

$E_{n+1} = E_{n+2} = \cdots = \varnothing$ にとれば，**1.4.6** の式は次のようになる．

$$P(E) \leqslant \sum_{\alpha=1}^{n} P(E_\alpha).$$

もちろんこれは $E, E_1, \cdots, E_n$ が $\mathscr{F}$ の集合でも成り立つ．

注意　標本空間 R が有限個の標本点のみ含むとき，初期事象族 $\mathscr{F}_0$ として，1 点からなる事象全体をとり，かつ $\mathscr{F}_0$ のあらゆる要素に 1 つの確率を割り当てれば，明らかにこの対応がブール確率測度を定める．これは，$\mathscr{F}_0$ で生成されるブール集合体 $\mathscr{F}$ のどの事象の確率も，$\mathscr{F}_0$ に含まれる集合に対応した確率から決定可能なことを意味している．

1.5　確率測度の拡張

(a)　確率測度の拡張の一意性

どのような初期ブール集合体 $\mathscr{F}$ から出発しても，$\mathscr{F}$ を含む最小のボレル集合体 $\mathscr{B}(\mathscr{F})$ が存在することが，1.3 節からわかっている．いまブール集合体 $\mathscr{F}$ 上の確率測度があるとする．$\mathscr{F}$ の集合に対して定義された確率にある演算を施しても，$\mathscr{F}$ の集合に対してすでに定めた確率はどれも変えないで，$\mathscr{B}(\mathscr{F})$ 上で定義された確率測度が得られるか，という疑問が出て来る．これは可能である．これは $\mathscr{F}$ 上の確率測度から $\mathscr{B}(\mathscr{F})$ 上の確率測度への拡張（または生成）とみなされる．

$\mathscr{F}$ 上の確率測度を $\mathscr{B}(\mathscr{F})$ 上へ拡張する方法があるかどうかを考える前に，その拡張の一意性に関する次の定理を述べておこう．

1.5.1　$\mathscr{F}$ はブール集合体，P_1，P_2 はボレル集合体 $\mathscr{B}(\mathscr{F})$ 上で定義された確率測度とする．$E \in \mathscr{F}$ なるすべての E に対して，$P_1(E) = P_2(E)$ ならば，すべての $E \in \mathscr{B}(\mathscr{F})$ に対しても $P_1(E) = P_2(E)$ である．

もし $\mathscr{H}$ を，任意の $E \in \mathscr{H}$ に対して $P_1(E) = P_2(E)$ となる $\mathscr{B}(\mathscr{F})$ に含まれる集合族とすれば，$\mathscr{H}$ はボレル集合体であることが証明できる．そして $\mathscr{H}$ は（最小の）ボレル集合体 $\mathscr{B}(\mathscr{F})$ を含むが，$\mathscr{H} \subset \mathscr{B}(\mathscr{F})$ であるから $\mathscr{H} = \mathscr{B}(\mathscr{F})$ となり，$\mathscr{B}(\mathscr{F})$ に含まれる任意の E に対して $P_1(E) = P_2(E)$ となる．

定理 **1.5.1** は，基本的にはブール集合体 $\mathscr{F}$ から生成されたボレル集合体上で定義され，$\mathscr{F}$ の集合では等しいが $\mathscr{B}(\mathscr{F}) - \mathscr{F}$ の集合では等しくない 2 つの確率測度はありえないことを述べたものである．このように，ブール集合体 $\mathscr{F}$ 上の確率測度から出発して，それをボレル集合体 $\mathscr{B}(\mathscr{F})$ へ拡張する方法を見い出せれば，その確率測度は **1.5.1** から一意的となる．

(b)　$\mathscr{F}$ から $\mathscr{B}(\mathscr{F})$ への確率測度の拡張

上で述べた拡張は Carathéodory (1927) の外測度の定理を用いている．[Halmos

(1950) をも見よ]．この問題に対して外測度を定義するには，まずブール集合体 $\mathscr{F}$ の集合に対する確率測度 P から始めよう．そこで $\mathscr{B}(\mathscr{F})$ の任意の集合 E をとり，$E_1, E_2, \cdots$ を $E \subset \bigcup_{\alpha=1}^{\infty} E_\alpha$，すなわち $\bigcup_{\alpha=1}^{\infty} E_\alpha$ は E の被覆となるような $\mathscr{F}$ の集合列とする．E の外測度 $P^*(E)$ は，$\mathscr{F}$ の集合で生成可能なすべての被覆に対して $\sum_{\alpha=1}^{\infty} P(E_\alpha)$ の最大の下界として定義される．すなわち

(1.5.1)
$$P^*(E) = \inf \sum_{\alpha=1}^{\infty} P(E_\alpha).$$

任意のブール集合体 $\mathscr{F}$ に対して，R の中の任意の集合 E の被覆として $E_1 = R$, $E_2 = E_3 = \cdots = \varnothing$ をとることができるから，R のいかなる集合に対しても $P^*(E)$ が存在することは明らかである．しかし実際に重要なのは，$\mathscr{B}(\mathscr{F})$ の集合 E に対する $P^*(E)$ だけである．

次の P^* の性質は読者各自で証明せよ．

(1.5.2)
$$P^*(\varnothing) = 0.$$

(1.5.3)
$$P^*(R) = 1.$$

(1.5.4)
$$E \in \mathscr{F} \quad ならば \quad P^*(E) = P(E).$$

(1.5.5)
$$E \subset F \quad ならば \quad P^*(E) \leqslant P(F).$$

(1.5.6)
$$E \subset \bigcup_{\alpha} E_\alpha \quad ならば \quad P^*(E) \leqslant \sum_{\alpha} P^*(E_\alpha).$$

ブール集合体 $\mathscr{F}$ 上の確率測度 P の，ボレル集合体 $\mathscr{B}(\mathscr{F})$ 上の確率測度への拡張に関する基本定理を次に述べる：

1.5.2　P をブール集合体 $\mathscr{F}$ 上の確率測度とする．そして (1.5.1) で定義した $P^*(E)$ は，すべての $E \in \mathscr{F}$ に対し $P^*(E) = P(E)$ となる $\mathscr{B}(\mathscr{F})$ 上の確率測度である．

証明は Halmos (1950) を参照せよ．

1.6　統計的独立性

(a)　標本空間の直積

標本空間 $R^{(1)}$ と $R^{(2)}$ の直積 R とは，順序のついた組 $(e^{(1)}, e^{(2)})$ 全体の集合である．

ただし $e^{(1)} \in R^{(1)}$, $e^{(2)} \in R^{(2)}$．これを次のように書く．

(1.6.1)
$$R = R^{(1)} \times R^{(2)}.$$

同様に $E^{(1)}$ と $E^{(2)}$ がそれぞれ $R^{(1)}$ と $R^{(2)}$ の事象であるとき，$E^{(1)}$ と $E^{(2)}$ の直積とは，その標本点 $e = (e^{(1)}, e^{(2)})$ が $e^{(1)} \in E^{(1)}$, $e^{(2)} \in E^{(2)}$ となる性質を持つ事象 E のことである．これを次のように書く．

(1.6.2)
$$E = E^{(1)} \times E^{(2)}.$$

E は $E^{(1)}$ または $E^{(2)}$ が空集合のとき，かつそのときに限って R において空集合となる．$E^{(1)}$ と $F^{(1)}$ が $R^{(1)}$ の事象で，$E^{(2)}$ と $F^{(2)}$ が $R^{(2)}$ の事象のとき，$(E^{(1)} \times E^{(2)}) \subset (F^{(1)} \times F^{(2)})$ となるための必要十分条件は $E^{(1)} \subset F^{(1)}$ かつ $E^{(2)} \subset F^{(2)}$．さらに $(E^{(1)} \times E^{(2)}) = (F^{(1)} \times F^{(2)})$ となるための必要十分条件は $E^{(1)} = F^{(1)}$ かつ $E^{(2)} = F^{(2)}$ である．

(1.6.2) で定義した E は，$E^{(1)}$ と $E^{(2)}$ の**結合生起**ということもある．また $E^{(1)}$ は E の $R^{(1)}$ 上への**射影**という．$E^{(2)}$ に対しても同様ないい方ができる．$R^{(1)}$ と $R^{(2)}$ は R の**成分**または**周辺標本空間**と呼ばれる．R の中で，$e^{(1)} \in E^{(1)}$ となるすべての点 $(e^{(1)}, e^{(2)})$ をとれば，R における**筒集合**が得られる．実際は直積 $E^{(1)} \times R^{(2)}$ である．同じく $R^{(1)} \times E^{(2)}$ は $e^{(2)} \in E^{(2)}$ となる筒集合である．このように，$E^{(1)} \times E^{(2)}$, $E^{(1)} \times R^{(2)}$, $R^{(1)} \times E^{(2)}$ は標本空間 $R = R^{(1)} \times R^{(2)}$ の事象で，直積 $E^{(1)} \times E^{(2)}$ が筒集合 $E^{(1)} \times R^{(2)}$ と $R^{(1)} \times E^{(2)}$ との共通部分となっている．すなわち

(1.6.3)
$$E^{(1)} \times E^{(2)} = (E^{(1)} \times R^{(2)}) \cap (R^{(1)} \times E^{(2)}).$$

直積の型をした R における事象 E は確率論において重要な役割を果たす特別な事象族であることを強調しておく．

直積の概念は各標本空間 $R^{(1)}, R^{(2)}, \cdots$ に含まれる有限または可付番無限個の事象 $E^{(1)}, E^{(2)}, \cdots$ へそのまま拡張できる．

例題　標本空間と事象の直積の一例として，$R^{(1)}$, $R^{(2)}$, $E^{(1)}$, $E^{(2)}$ が実数 R_1 の軸上の点集合の場合を考えてみる．図 1.2 で示したように $R^{(1)}$ は $e^{(1)}$ 軸上の閉区間 $[a_1, b_1]$，$R^{(2)}$ は $e^{(2)}$ 軸上の閉区間 $[a_2, b_2]$，$E^{(1)}$ は $R^{(1)}$ の中の閉集合，$E^{(2)}$ は $R^{(2)}$ の中の閉集合とする．直積空間 $R = R^{(1)} \times R^{(2)}$ はその境界も含めた大きい方の矩形 $ABCD$ の中のすべての点からなっている．筒集合 $E^{(1)} \times R^{(2)}$ は傾線を入れた垂直の細長い部分に含まれる点集合からなり，また筒集合 $R^{(1)} \times E^{(2)}$ は傾線を入れた水平の細長い部分に含まれる点集合からなっている．そして直積 $E = E^{(1)} \times E^{(2)}$ は 4 つの黒い矩

形の点集合からなる．このように E は R 内に示された 2 つの筒集合の共通部分である．すなわち $E=(E^{(1)}\times R^{(2)})\cap(R^{(1)}\times E^{(2)})$.

事象の直積をつくる演算は実数で表わされる事象，すなわち直線や平面または多次元ユークリッド空間内の点集合で表わされる事象より，さらに一般的な事象に適用できることを強調しておく．たとえば，トランプで 2 回勝負する場合，$e^{(1)}$ を競技者 A が最初の勝負のときに手にした持札とし，$e^{(2)}$ を 2 回目の持札とする．それぞれ事象空間 $R^{(1)}$, $R^{(2)}$ に含まれる標本点 $e^{(1)}$ と $e^{(2)}$ の数は $\binom{52}{13}$ である．さらに，A がその 2 回の勝負で手にできる $\binom{52}{13}^2$ 個の起りうる持札の組 $(e^{(1)},e^{(2)})$ は直積空間 $R=R^{(1)}\times R^{(2)}$ の $\binom{52}{13}^2$ 個の標本点に相当する．いま $E^{(1)}$ が 2 枚のエースを含んだすべての持札からなる $R^{(1)}$ の部分集合を示し，$E^{(2)}$ が 3 枚のキングを含むすべての持札からなる $R^{(2)}$ の部分集合を示すとすれば，$E=E^{(1)}\times E^{(2)}$ は，最初の持札 $e^{(1)}$ が 2 枚のエースを含み，2 回目の持札 $e^{(2)}$ が 3 枚のキングを含む持札の組 $(e^{(1)},e^{(2)})$ のすべてを含むような R の事象である．このような持札の組すなわち E の標本点は $\left[\binom{4}{2}\binom{48}{11}\right]\left[\binom{4}{3}\binom{48}{10}\right]$ 個ある．

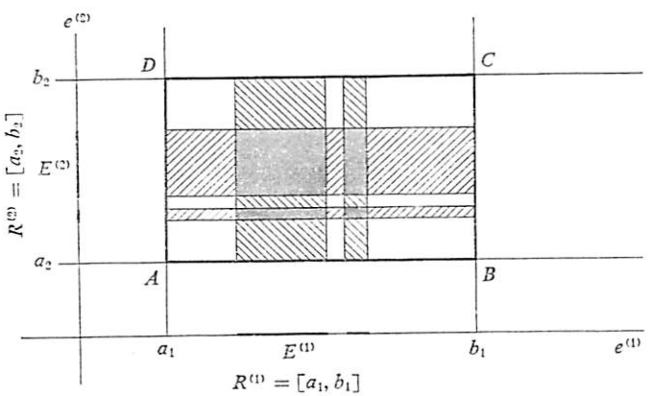

図 1.2　事象および標本空間の直積の図．

$R^{(1)}$ と $R^{(2)}$ を 2 つの標本空間とし，$\mathscr{B}(\mathscr{F}^{(1)})$ と $\mathscr{B}(\mathscr{F}^{(2)})$ をそれぞれこれら 2 つの標本空間のブール集合体 $\mathscr{F}^{(1)}$ および $\mathscr{F}^{(2)}$ で生成されるボレル集合体とする．ここで，$E^{(1)}\in\mathscr{B}(\mathscr{F}^{(1)})$, $E^{(2)}\in\mathscr{B}(\mathscr{F}^{(2)})$ であるような $E=E^{(1)}\times E^{(2)}$ の型をした $R=R^{(1)}\times R^{(2)}$ の集合族 $\mathscr{F}_0$ を考えてみる．初期族として $\mathscr{F}_0$ をとれば，R における最小のボレル集合体を生成することができる．それを次のように表わす．

(1.6.4)
$$\mathscr{B}(\mathscr{F}_0).$$

実際には，このボレル集合体は初期族として $E^{(1)}\times E^{(2)}$ の型をした集合をとっても生成できることが証明できる．ただし $E^{(1)}\in\mathscr{F}^{(1)}$ かつ $E^{(2)}\in\mathscr{F}^{(2)}$ である．

(b) 確率測度の積

ここで $P^{(1)}$ および $P^{(2)}$ をそれぞれ $\mathscr{B}(\mathscr{F}^{(1)})$, $\mathscr{B}(\mathscr{F}^{(2)})$ 上の確率測度とする．$\mathscr{F}_0$ の任意の集合 $E=E^{(1)}\times E^{(2)}$ に対して，E の確率を次式で定める．ただし $E^{(1)}\in\mathscr{B}(\mathscr{F}^{(1)})$ かつ $E^{(2)}\in\mathscr{B}(\mathscr{F}^{(2)})$ である．

(1.6.5)
$$P(E)=P^{(1)}(E^{(1)})\cdot P^{(2)}(E^{(2)}).$$

P は $\mathscr{F}_0$ 上のブール確率測度であり，$\mathscr{B}(\mathscr{F}_0)$ へ一意的に拡張できることが示せる．

ボレル集合体 $\mathscr{B}(\mathscr{F}_0)$ 上の確率測度 P が，$E^{(1)}\in\mathscr{B}(\mathscr{F}^{(1)})$, $E^{(2)}\in\mathscr{B}(\mathscr{F}^{(2)})$ となるすべての集合 $E=E^{(1)}\times E^{(2)}$ に対して，(1.6.5) を満足するとき，**成分確率空間** $(R^{(1)},\mathscr{B}(\mathscr{F}^{(1)}),P^{(1)})$ と $(R^{(2)},\mathscr{B}(\mathscr{F}^{(2)}),P^{(2)})$ は**（統計的に）独立**であるという．確率空間

(1.6.6)
$$(R,\mathscr{B}(\mathscr{F}_0),P)$$

は P が (1.6.5) で定義された確率測度のとき，下記の 2 つの成分確率空間の直積と呼ばれる．

$$(R^{(1)},\mathscr{B}(\mathscr{F}^{(1)}),P^{(1)}),\quad (R^{(2)},\mathscr{B}(\mathscr{F}^{(2)}),P^{(2)}).$$

また，統計的独立性の概念は 2 つ以上の確率空間にも拡張できる．

1.7 確率変数

$(R,\mathscr{B},P)$ を確率空間とする．$x(e)$ は R のすべての標本点 e で定義された 1 価の実関数で，任意の実数 b に対して $x(e)\leqslant b$ となる R の事象 E_b は $\mathscr{B}$ に属するものとする．このとき関数 $x(e)$ を $\mathscr{B}$ に関する**確率変数**と呼ぶ．$x(e)$ は $\mathscr{B}$ 可測であるともいう．$\{E_b\}$ ですべての実数 b に対応する R の事象族を表わせば，$\mathscr{B}(\{E_b\})$ はこの族で生成されるボレル集合体である．すべての $e\in R$ に対して $x(e)$ がとる数の集合を $x(e)$ の**標本空間**と呼ぶ．

このように，$\mathscr{B}_1$ が実直線 R_1 上のボレル集合族ならば，確率変数 $x(e)$ はすべてのボレル集合 $E'\in\mathscr{B}_1$ に対して，$x(e)\in E'$ となる R の標本点すべてからなる事象 $E\in\mathscr{B}$ があるように，R の標本点 e を R_1 の標本点 x に写像する．便宜上，E を $x^{-1}(E')$ で

表わす．$P'(E') = P(x^{-1}(E'))$ とおけば，基本標本空間 R に関する確率空間 $(R, \mathscr{B}, P)$ から実直線 R_1 に関する確率空間 $(R_1, \mathscr{B}_1, P')$ を定義することができる．このようにして $(R_1, \mathscr{B}_1, P')$ は確率変数 $x(e)$ によって $(R, \mathscr{B}, P)$ から導かれるという．$(R_1, \mathscr{B}_1, P')$ が確率空間であるという証明は練習問題として読者に残しておく．

与えられた確率空間 $(R, \mathscr{B}, P)$ に対して，$x_1(e), \cdots, x_k(e)$ を k 個の確率変数とすれば，$b_1, \cdots, b_k$ を任意の実数として，

$$\{e : x_i(e) \leqslant b_i,\ i = 1, \cdots, k\}$$

の型をした R の事象 $E_{b_1, \cdots, b_k}$ はやはり $\mathscr{B}$ に属し，$P(E_{b_1, \cdots, b_k}) = P(x_1(e) \leqslant b_1, \cdots, x_k(e) \leqslant b_k)$ となる．さらに E' を $\mathscr{B}_k$（k 次元ユークリッド空間 R_k のボレル集合）の任意の集合とすれば，$(x_1(e), \cdots, x_k(e)) \in E'$ となるような点 e の集合 E もまた $\mathscr{B}$ に含まれる．したがって $P'(E') = P(E)$ と定めれば，$(x_1(e), \cdots, x_k(e))$ によって $(R, \mathscr{B}, P)$ から導かれる確率空間 $(R_k, \mathscr{B}_k, P')$ が得られる．$(x_1(e), \cdots, x_k(e))$ を $\mathscr{B}$ に関する **k 次元またはベクトル確率変数**または $\mathscr{B}$ 可測 k 次元確率変数という．

R における任意の標本点 e が $M < +\infty$ に対して $|x(e)| < M$ となるならば，$x(e)$ を**有界確率変数**と呼ぶ．k 次元確率変数は各成分が有界であれば有界である．

例題　例題で確率変数の概念を明確にしよう．

一組のトランプを 13 枚配る場合を考えよう．基本標本空間 R の標本点 e の集合は $\binom{52}{13}$ 通りの起りうる場合からなっている．$x(e)$ を e に含まれるエースの枚数をとる確率変数とすれば，$x(e)$ は R におけるすべての標本点で定義されていて，R の点で値 0，1，2，3，4 のどれかをとる．このように $x(e)$ は R のあらゆる標本点 e を実直線 R_1 上の点 0，1，2，3，4 のどれか 1 つに写像する．R における標本点の数は有限であるから，R におけるすべての起りうる事象の族（R および空集合を含む）はブール集合体 $\mathscr{F}$ をなす．個々の実数 b に対して $x(e) \leqslant b$ となる R の事象は族 $\mathscr{F}$ に含まれる事象の 1 つである．R の中の個々の標本点に確率を定めれば（札が“完全に”切ってあれば，確率はすべて $1/\binom{52}{13}$ になる），すべての実数 b に対して事象 $E_b = \{e ; x(e) \leqslant b\}$ の確率を定めるのに十分なブール確率空間 $(R, \mathscr{F}, P)$ が得られる．このように E' を R_1 の任意の（ボレル）集合とするとき，その確率は $x(e)$ によって $(R, \mathscr{F}, P)$ から導かれる確率空間 $(R_1, \mathscr{B}_1, P')$ で与えられる．ここで $\mathscr{B}_1$ で必要なものは 5 点からなる集合 $\{0, 1, 2, 3, 4\}$ とその部分集合，すなわち $x(e)$ の標本空間とその部分集合である．

$(x_1(e), x_2(e))$ が 2 次元の確率変数で，$x_1(e)$ を標本点 e に含まれるエースの枚数，$x_2(e)$ をスペードの枚数とすれば，確率 $P(x_1(e) \leqslant b_1, x_2(e) \leqslant b_2)$ がブール確率空間 $(R, \mathscr{F}, P)$ から計算できるような 2 次元確率変数が得られたことになる．すべての実数の組 (b_1, b_2) に対して区間 $(-\infty, -\infty ; b_1, b_2]$ に上記の確率を定めれば，確率空間 $(R_2, \mathscr{B}_2, P')$ がベクトル確率変数 $(x_1(e), x_2(e))$ により導かれ，R_2 のあらゆる事象 E の確率が与えられる．ここで，R_2 の事象 E で実際に必要なのは 62 点からなる集合 $\{(x_1, x_2) : x_1 = 0, 1, 2, 3, 4 ; x_2 = 0, 1, \cdots, 13,$ ただし $0 \leqslant x_1 + x_2 \leqslant 14$ しかし $(4, 0)$ と $(0, 13)$ はのぞく$\}$ に含まれる集合，すなわち $(x_1(e), x_2(e))$ の標本空間とその部分集合である．

もう 1 つ別の例を考えてみよう．2 個の電球 B_1，B_2 の点灯時間を考えよう．そしてそれらの寿命をそれぞれ t_1 と t_2 とする．基本標本空間 R は t_1t_2 平面の第 1 象限であり，標本点 e はその象限内の点である．どちらの電球も一方の電球よりある時間単位 b 以上に寿命の長くない事象 E_b は，$|t_1 - t_2| \leqslant b$ なる R の点すべてから成り立っている．このように R における各点 $e = (t_1, t_2)$ に $|t_1 - t_2|$ で確率変数 $x(e)$ を定義できる．あらゆる実数 b に対して $|t_1 - t_2| \leqslant b$ となる事象を含むボレル集合体 $\mathscr{B}$ と，$\mathscr{B}$ 上の確率測度 P に関する確率空間 $(R, \mathscr{B}, P)$ が与えられれば $x(e) \leqslant b$ となる確率は計算できる．確率変数 $|t_1 - t_2|$ は R の点を実直線 R_1 に写像し，b を任意の実数，E'_b を R_1 における区間 $(-\infty, b]$，E_b を $|t_1 - t_2| \leqslant b$ となる R の集合としたとき，$P'(E'_b) = P(E_b)$ となる．このように，初期事象族 $\{E_b : b$ は実数$\}$ が得られ，それは確率 $\{P(E_b)\}$ と共に最小のボレル集合体 $\mathscr{B}(\{E_b\})$ を生成し，$x(e)$ によって確率空間 $(R_1, \mathscr{B}_1, P')$ が導かれるような確率空間 $(R, \mathscr{B}(\{E_b\}), P)$ を生成する．そして $|t_1 - t_2|$ が $\mathscr{B}_1$ のどの集合 E' に属する場合に対しても確率が与えられる．

$x_1(e) = t_1 + t_2$，$x_2(e) = t_1 - t_2$ とすれば，R のあらゆる点で定義された 2 つの確率変数が得られる．ここで，$E_{b_1 b_2} = \{e : x_1(e) \leqslant b_1,\ x_2(e) \leqslant b_2\}$ である初期事象族 $\{E_{b_1 b_2} : b_1, b_2$ は実数$\}$ と，その確率 $\{P(E_{b_1 b_2})\}$ から，確率空間 $(R, \mathscr{B}(\{E_{b_1 b_2}\}), P)$ が生成されれば，$(x_1(e), x_2(e))$ は $(R, \mathscr{B}(\{E_{b_1 b_2}\}), P)$ から $(R_2, \mathscr{B}_2, P')$ を導き，それは $(t_1 + t_2, t_1 - t_2)$ が $\mathscr{B}_2$ のどの集合 E' に属する場合にも確率を与える．

1.8　確率変数の積分

本節では，確率論の用語を説明し，また証明なしに基本的なルベーグ=スティルチェス積分論をいくらか述べてみる．証明は Halmos (1950)，Loève (1955)，McShane と Botts (1959)，Saks (1937) などを参照されたい．

$(R, \mathscr{B}, P)$ を確率空間，$x(e)$ を $\mathscr{B}$ に関する確率変数とする．$x(e)$ がすべての $e \in$

E_i, $i=1, \cdots, k$ に対して，$x(e)=x_i$ となる有限個の相異なる値 $x_1, \cdots, x_k$ しかとらないとき，$x(e)$ は単純確率変数*) と呼ばれる．そして次のように書く．

(1.8.1)
$$\int_R x(e)\,dP(e)=\sum_{i=1}^{k} x_i P(E_i).$$

さらに一般的に，E が $\mathscr{B}$ における任意の集合の場合には次のように書く．

(1.8.2)
$$\int_E x(e)\,dP(e)=\sum_{i=1}^{k} x_i P(E\cap E_i).$$

いま，$x(e)$ が有界で，かつ無限に多くの値をとれる場合を考えてみる．このとき，次のような単純確率変数 $x_1(e), x_2(e), \cdots$ の列が，すべての $e\in R$ に対して一様に存在することが証明される．

(1.8.3)
$$\lim_{\alpha\to\infty} x_\alpha(e)=x(e).$$

さらに個々の $E\in\mathscr{B}$ に対して

(1.8.4)
$$\lim_{\alpha\to\infty}\int_E x_\alpha(e)\,dP(e)$$

が存在し，この極限は $x_1(e), x_2(e), \cdots$ の選び方にはよらない．この極限を

(1.8.5)
$$\int_E x(e)\,dP(e)$$

と表わし，E 上の $(R,\mathscr{B},P)$ に関する $x(e)$ のルベーグ=スティルチェス積分と呼ぶ．単純確率変数 $x(e)$ の場合には E 上のルベーグ=スティルチェス積分は (1.8.2) によって与えられる．さらに一般的には $x(e)$ が有界である必要はないが，非負のときはすべての $e\in R$ に対して次の 2 つの式を満たす有界な確率変数の列 $x_1(e)\leqslant x_2(e)\leqslant\cdots$ が存在すれば積分可能である．

(1.8.6)
$$\lim_{\alpha\to\infty} x_\alpha(e)=x(e)\qquad (\text{すべての } e\in R \text{ に対して}),$$

(1.8.7)
$$\lim_{\alpha\to\infty}\int_E x_\alpha(e)\,dP(e)<+\infty.$$

この極限が $(R,\mathscr{B},P)$ に関する E 上の $x(e)$ の積分であり，次のように表わされる．

$$\int_E x(e)\,dP(e).$$

これらの条件のもとで，任意の $E\in\mathscr{B}$ に対し

*) 原書では "simple" である．ベーグ積分論では，通常，階段関数と呼ばれている．あるいは単関数ともいわれる．（訳注）

(1.8.8)
$$\lim_{\alpha\to\infty}\int_E x_\alpha(e)\,dP(e)$$

が存在し，(1.8.6)，(1.8.7) を満足するように選んだ特別な列 $x_1(e)\leqslant x_2(e)\leqslant\cdots$ とは無関係であることが証明できる．

最後に，任意の $x(e)$ に対しては，$x'(e)$ と $x''(e)$ が非負の積分可能な確率変数でその差 $x'(e)-x''(e)$ として表わされるとき，積分可能であるといい，この場合，次のように定義する．

(1.8.9)
$$\int_E x(e)\,dP(e)=\int_E x'(e)\,dP(e)-\int_E x''(e)\,dP(e).$$

(1.8.9) の左辺の値は $x'(e)$ と $x''(e)$ の選び方には無関係であることが証明できる．

ルベーグ=スティルチェス積分は次のような重要な性質を持っている．ただし $x(e)$, $x_1(e), x_2(e), \cdots$ は確率変数で，E は $\mathscr{B}$ の任意の集合である．

1.8.1　すべての $e\in R$ に対して，$x(e)=k$ ならば，k は定数

$$\int_E x(e)\,dP(e)=kP(E).$$

1.8.2　a と b は定数で，すべての $e\in R$ に対して $a\leqslant x(e)\leqslant b$ ならば

$$aP(E)\leqslant\int_E x(e)\,dP(e)\leqslant bP(E).$$

1.8.3　$x_1(e)$ と $x_2(e)$ が積分可能で，a と b が任意の定数であれば，$ax_1(e)+bx_2(e)$ は積分可能で，次のようになる．

$$\int_E (ax_1(e)+bx_2(e))\,dP(e)=a\int_E x_1(e)\,dP(e)+b\int_E x_2(e)\,dP(e).$$

1.8.4　$x_1(e)$，$x(e)$，$x_2(e)$ が積分可能で，すべての $e\in R$ に対して $x_1(e)\leqslant x(e)\leqslant x_2(e)$ ならば

$$\int_E x_1(e)\,dP(e)\leqslant\int_E x(e)\,dP(e)\leqslant\int_E x_2(e)\,dP(e).$$

1.8.5　$x(e)$ が積分可能なための必要十分条件は $|x(e)|$ が積分可能な場合である．さらに

$$\left|\int_E x(e)\,dP(e)\right|\leqslant\int_E |x(e)|\,dP(e).$$

1.8.6　$x_1(e), x_2(e), \cdots$ を確率変数列とし，$y(e)$ が積分可能で，すべての $e \in R$ に対して

$$\lim_{\alpha\to\infty} x_\alpha(e) = x(e) \quad \text{かつ} \quad |x_\alpha(e)| \leqq y(e)$$

ならば，$x(e)$ は積分可能で

$$\lim_{\alpha\to\infty} \int_E |x_\alpha(e) - x(e)|\, dP(e) = 0.$$

特にすべての $E \in \mathscr{B}$ に対して一様に

$$\lim_{\alpha\to\infty} \int_E x_\alpha(e)\, dP(e) = \int_E x(e)\, dP(e)$$

となる．

1.8.7　$(R_k, \mathscr{B}_k, P')$ をベクトル確率変数 $(x_1(e), \cdots, x_k(e))$ によって確率空間 $(R, \mathscr{B}, P)$ から導かれる確率空間とする．$g(x_1, \cdots, x_k)$ を $\mathscr{B}_k$ 可測とする（したがって $g(x_1(e), \cdots, x_k(e))$ は $\mathscr{B}$ に関して可測である）．E を $(x_1(e), \cdots, x_k(e)) \in E'_k$ となる R の集合とすれば

$$\int_E g(x_1(e), \cdots, x_k(e))\, dP(e) = \int_{E'_k} g(x_1, \cdots, x_k)\, dP'(x_1, \cdots, x_k).$$

一方の積分が有限ならば他方も有限で，2 つは等しいという意味である．ただし $E'_k \in \mathscr{B}_k$ である．

最後に **1.8.1** から **1.8.6** まで “すべての $e \in R$” という個所を “$P(F) = 0$ となる集合 F をのぞいたすべての $e \in R$” と置き換えても結果はそのままで変わらない．

1.9　条件つき確率

$(R, \mathscr{B}, P)$ を確率空間，E_1，E_2 を $P(E_1) > 0$ となる $\mathscr{B}$ の事象とする．

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} \tag{1.9.1}$$

と書き，この比を E_1 が起ったという条件の下での事象 E_2 の**条件つき確率**と呼ぶ．

$P(E_1) > 0$ となる $\mathscr{B}$ に含まれる任意の E_1 を固定すれば，$(R, \mathscr{B}, P(\cdot\,|E_1))$ は確率空間となることが証明できる．ここで $P(\cdot\,|E_1)$ は $E_2 \in \mathscr{B}$ 上で値 $P(E_2|E_1)$ をとる測度である．

(1.9.1) は次のように書き直せる．

$$P(E_1 \cap E_2) = P(E_1)\cdot P(E_2|E_1). \tag{1.9.2}$$

$P(E_2|E_1) = P(E_2)$ のときは，E_1 と E_2 は独立事象となり，次のようになる．

$$P(E_1 \cap E_2) = P(E_1)P(E_2). \tag{1.9.2 a}$$

もっと一般的には次のようにいえる．

1.9.1　$E_1, E_2, \cdots$ が $\mathscr{B}$ における有限または可付番無限列で，$P(E_1) > 0$，$P(E_1 \cap E_2) > 0$，$P(E_1 \cap E_2 \cap E_3) > 0$，$\cdots$ となる確率測度 P を持つならば

$$P(E_1 \cap E_2 \cap \cdots) = P(E_1)\cdot P(E_2|E_1)\cdot P(E_3|E_1 \cap E_2)\cdots \cdot P(E_n|E_1 \cap E_2 \cap \cdots \cap E_{n-1})\cdots. \tag{1.9.3}$$

$E_1, E_2, \cdots$ が互いに独立な場合には (1.9.3) は次のようになる．

$$P(E_1 \cap E_2 \cap E_3 \cap \cdots) = P(E_1)\cdot P(E_2)\cdot P(E_3)\cdots. \tag{1.9.3 a}$$

もう 1 つ重要な結果を次に示す．

1.9.2　$E_1, E_2, \cdots$ が $\mathscr{B}$ における有限または可付番無限個の互いに素な列で，0 でない確率を持ち，かつ $\bigcup_\alpha E_\alpha = R$ のとき，E を $\mathscr{B}$ における任意の集合とすれば，

$$P(E) = P(E_1)P(E|E_1) + P(E_2)P(E|E_2) + \cdots. \tag{1.9.4}$$

この証明は読者にゆだねる．

1.10　条件つき確率変数

与えられた確率空間 $(R, \mathscr{B}, P)$ に対して，$(x_1(e), x_2(e))$ を $\mathscr{B}$ に関する 2 次元の確率変数とする．(1.9.1) で述べた集合 E_1 および E_2 を，E'_1，E'_2 が $\mathscr{B}_1$ の集合で，それぞれ $x_1(e) \in E'_1$，$x_2(e) \in E'_2$ となる集合として選べば，(1.9.1) は確率変数に関する条件つき確率の公式となる．そしてそれは $x_1(e) \in E'_1$ が与えられたときの $x_2(e) \in E'_2$ の条件つき確率を示すものである．特に，E'_1 が 1 点 x_1 のみであると仮定する．$P(x_1(e) = x_1) > 0$ ならば，(1.9.1) について，何も支障はない．しかし，$P(x_1(e) = x_1) = 0$ のときは次のような疑問が出る：条件つき確率 $P(x_2(e) \in E'_2|x_1(e) = x_1)$ に

意味を持たせることに意義があるか．

数理統計学では，この条件つき確率に対してかなり本質的な考察をして意味を持たせられる．これは 2.9 節に示す．しかしさらに一般的な条件の下では，この質問に対する答はラドン=ニコディムの定理で与えられる．それを次に示す．

1.10.1　$(R, \mathscr{B}, P)$ が確率空間で，Q が $\mathscr{B}$ 上の（有限）完全加法的集合関数かつ $P(E)=0$ となるあらゆる集合 $E \in \mathscr{B}$ に対して $Q(E)=0$ となるならば，次のような確率変数 $g(e)$ が存在して，これがすべての $E \in \mathscr{B}$ に対して成り立っている．

$$Q(E) = \int_E g(e)\, dP. \tag{1.10.1}$$

さらに このような 2 つの確率変数 $g(e)$ と $h(e)$ に対して $P(g(e) \fallingdotseq h(e)) = 0$.

この定理の証明は省略する．さらに一般化した公式およびその証明は Halmos (1950) の本を参照されたい．

ここで特に注意しなければならないのは，Q の値に非負という制約がないことである．

$\mathscr{B}$ 上で定義された 2 つの完全加法的集合関数 P および Q が $P(E)=0$ となるすべての集合 E に対して $Q(E)=0$ ならば，Q は P に関して絶対連続であるという．

$(R, \mathscr{B}, P)$ を確率空間とし，$x(e)$ を確率変数とする．与えられた $E \in \mathscr{B}$ と一定の x_1 に対して条件つき確率 $P(E|x(e)=x_1)$ を見つけたい．定義 (1.9.1) は $P(x(e)=x_1) = 0$ の場合には意味がなくなる．

しかし直観的に，この条件つき確率を，$P(E|x(e) \in N_\alpha(x_1))$ の $\alpha \longrightarrow \infty$ としたときのある種の“極限”として定義できそうな気がする．ただしここで $N_1(x_1), N_2(x_1), \cdots$ は点 x_1 に収束する x_1 の近傍列とする．一般にこの極限はすべて存在するとは限らない．しかし，$F \in \mathscr{B}_1$ なる適当な領域内で確かに存在するならば，あらゆる $x_1 \in F$ 上で適当に $P(E|x(e)=x_1)$ を平均化することによって，$P(E|x(e) \in F)$ が得られることは可能に思える．

一般的に，上で述べたような極限を特殊な条件以外で定めることは困難である．しかし Kolmogorov (1933 *a*) は次のような方法を提唱した．

任意の $E \in \mathscr{B}$，$F \in \mathscr{B}_1$ に対して

$$Q(F) = P(E \cap x^{-1}(F)) \tag{1.10.2}$$

とする．$P(x^{-1}(F))=0$ ならば $Q(F)=0$ であり，Q は $\mathscr{B}_1$ 上の完全加法的集合関数であることは明らかである．**1.10.1** から，次のような実数値をとる $\mathscr{B}_1$ 可測な関数 $f(x)$ が存在する．

$$Q(F) = \int_F f(x)\, dP'(x). \tag{1.10.3}$$

ここですべての集合 $F \in \mathscr{B}_1$ に対し，$P'(F) = P(x^{-1}(F))$ である．したがって $g(e) = f(x(e))$ と書けば

$$P(E \cap x^{-1}(F)) = \int_{x^{-1}(F)} g(e)\, dP(e) \tag{1.10.4}$$

となり，$g(e)$ すなわち $f(x(e))$ は“与えられた $x(e)$ の下での E の条件つき確率”である．$f(x(e))$ は **1.10.1** の意味で一意的である．この“与えられた $x(e)$ の下での E の条件つき確率”の定義から，次のことが証明できる．R_1 における確率 0 の集合をのぞいて，

$$f(x_1) = \lim_{\alpha \to \infty} P(E|x(e) \in N_\alpha(x_1)) \tag{1.10.5}$$

が得られる．ただし $\{N_\alpha(x_1),\ \alpha = 1, 2, \cdots\}$ は x_1 に収束する x_1 の可測な近傍の列である．

この方法はそのままでベクトル確率変数 $(x_1(e), \cdots, x_k(e))$ の場合にも拡張できる．

問　　題

1.1　N 個の異なった名前の表から 2 つの名前を無作為に選び，アルファベット順に並べる．この操作からつくられる標本空間を述べ，それがいくつの標本点を含むかをいえ．またこれを N 個の名前の表から“無作為に” n 個の名前を抜き取る場合についても一般化せよ．

1.2　紳士録から“無作為に”選んだ 2 人の人間 A，B の誕生日（月と日）を記録する．うるう年を無視して，この操作でつくられる標本空間を述べ，その中に含まれる標本点の数を述べよ．また次の事象に含まれる標本点の数を示せ．

(a)　“A と B はまったく同じ誕生日を持つ．”

(b)　“A と B の誕生日は r 日以上離れていない．”

(c)　“A と B の誕生日は月が違っている．”

1.3　ある店の開店を 9 時，閉店を 17 時とする．“無作為に”選んだ買物客が歩いて

来て，x 時に店にはいり，y 時に出ていく（x も y も 9 時で始まる時間の軸上で時間の単位で測ったものとする）．このとき，(x, y) の標本空間を記せ．また次の事象を x と y を用いて述べよ：

（a）“買物客はその店に 1 時間以内しかいない．”

（b）“z 時に買物客はその店の中にいる．”

（c）“買物客はその店に u 時より前にはいり，v 時以後に出て行った．”

1.4 N 個の電球がはいっている箱に，フィラメントが切れた電球が r 個（$r < N$）混っている．ある人が電球を 1 つ 1 つ，欠陥電球（フィラメントが切れた電球）が 1 個見つかるまで調べるとする．この操作でつくられる標本空間を記せ．その標本空間にはいくつの点が含まれるか？　これを s 個の欠陥電球が見つかるまでに一般化せよ．

1.5 R を大学 A の全学生の集合とする．E_1 は雑誌 M_1 をとっている A の学生の集合を表わし，E_2 は雑誌 M_2 をとっている学生，E_3 は雑誌 M_3 をとっている学生の集合を表わすものとする．

（a）次の集合を言葉で述べよ：

$$E_1 \cup E_2 \cup E_3; \quad E_1 \cap E_2 \cap E_3; \quad R - (E_1 \cup E_2);$$
$$\overline{E_1} \cup \overline{E_2} \cup \overline{E_3}; \quad E_1 \cap E_2 \cap \overline{E_3}; \quad \overline{E_1} \cap \overline{E_2} \cap \overline{E_3}.$$

（b）次の集合を R, E_1, E_2, E_3 上での演算で表わせ．

（i）3 冊の雑誌のうちの 2 冊またはそれ以上をとっている学生．

（ii）3 冊の雑誌のうちの 1 冊より多くはとっていない学生．

1.6 一組のトランプから 13 枚配り，起り得る持札全体の集合を R とする．E_1, E_2, E_3, E_4 はそれぞれスペードのエース，ハートのエース，ダイヤのエース，クラブのエースを含んでいる 13 枚の異なった持札の集合とする．次の集合を言葉で述べ，おのおの集合には何通りの持札があるかをいえ．

$$E_1 \cap E_2; \quad E_1 - (E_2 \cap E_3); \quad (E_1 \cap E_2 \cap E_3) \cup E_4; \quad \overline{E_1} \cup \overline{E_2};$$
$$E_1 \cup (E_2 \cap E_3); \quad (E_1 \cup E_2) - \overline{E_3}; \quad \overline{E_1} \cap (\overline{E_2} \cup \overline{E_3});$$
$$(E_1 \cup E_2 \cup E_3) - E_4; \quad R - [(E_1 \cup E_2) \cap (E_3 \cup \overline{E_4})].$$

1.7 クラップゲーム*)は普通の 6 面のサイコロ 2 個を使って次のように行なう．投げる者は 7 または 11 を出せば，それで勝ちとなり，もし 2, 3 または 12 が出れば，それで負けとなる．4, 5, 6, 8, 9 または 10 が出た場合は，7 かまたは最初に投げたときの数が出るまで続けて投げる．最初に 7 が現われると負けである．また最初に投げたときの点が出れば勝ちとなる．このときの標本空間を述べ，実際，完全なサイコロを使った場合，勝つ確率は 244/495 であることを示せ．

1.8 標本空間 R が N 個の標本点を含むとき，これらの点で生成されるブール集合体に含まれる事象の数は 2^N 個であることを示せ．

1.9 xy 平面での 2 つの無限集合列 $E_1, E_2, \cdots$ と $F_1, F_2, \cdots$ を考える．ここで E_n は

*) 一種のばくちで，2 個のサイコロを振って出た目の合計で勝負を競う．（訳注）

$x^2 + y^2 < (1+n)/n$ となる点集合であり，F_n は $x^2 + y^2 \leqslant n/(1+n)$ となる点集合である．$\overline{E_n}$ は xy 平面全体に関する E_n の補集合，$\overline{F_n}$ も同様に定義されたものとする．次の集合を x と y を用いて述べよ．

$$\lim_\alpha E_\alpha; \quad \lim_\alpha \overline{E_\alpha}; \quad \lim_\alpha F_\alpha; \quad \lim_\alpha \overline{F_\alpha}; \quad \lim_\alpha (E_\alpha \cap \overline{F_\alpha}).$$

1.10 $E_1, \cdots, E_n$ を標本空間 R における任意の事象，$\overline{E_1}, \cdots, \overline{E_n}$ をその補集合とするとき，

$$\bigcup_{\alpha=1}^{n} E_\alpha \quad と \quad \bigcap_{\alpha=1}^{n} \overline{E_\alpha}$$

は互いに素であること，およびこの 2 つの合併集合は R であることを証明せよ．またそのことから

$$P\left(\bigcup_{\alpha=1}^{n} E_\alpha\right) = 1 - P\left(\bigcap_{\alpha=1}^{n} \overline{E_\alpha}\right)$$

となることを示せ．

1.11 事象 $E_1, \cdots, E_n$ の任意の指定された E_r が起る確率は P_r, $r = 1, \cdots, n$ である．このとき次のことを証明せよ．

（a）事象 $E_1, \cdots, E_n$ のうちの 1 つまたはそれ以上が起る確率は

$$\binom{n}{1} p_1 - \binom{n}{2} p_2 + \cdots + (-1)^{n-1} \binom{n}{n} p_n.$$

（b）事象 $E_1, \cdots, E_n$ のうちの m 個またはそれ以上が起る確率は

$$\binom{m-1}{m-1}\binom{n}{m} p_m - \binom{m}{m-1}\binom{n}{m+1} p_{m+1} + \cdots + (-1)^{n-m} \binom{n-1}{m-1}\binom{n}{n} p_n.$$

（c）事象 $E_1, \cdots, E_n$ のうちのちょうど m 個が起る確率は

$$\binom{m}{m}\binom{n}{m} p_m - \binom{m+1}{m}\binom{n}{m+1} p_{m+1} + \cdots + (-1)^{n-m} \binom{n}{m}\binom{n}{n} p_n.$$

1.12 コーンフレークを製造しているある会社では，包装したものに $1, 2, \cdots, r$ と番号を書いたカードを無作為につけることにしている．番号は全部同じ枚数だけ書いてある．n 個の（$n > r$）コーンフレークを買ったとき，それらの包から少なくとも完全な一組のカードを集められる確率は下記の通りであることを示せ．

$$1 - \binom{r}{1}\left(1 - \frac{1}{r}\right)^n + \binom{r}{2}\left(1 - \frac{2}{r}\right)^n + \cdots + (-1)^{r-1}\binom{r}{r-1}\left(1 - \frac{r-1}{r}\right)^n.$$

1.13 壺に $1, 2, \cdots, N$ の番号をつけた N 個の札がはいっていて，2 枚の札を（戻さないで）続けて取り出す．この操作で生じる標本空間 R の任意の点を e とする．$x(e)$ で，e を生じた 2 枚の札の番号の差の絶対値を表わす．標本空間 R のすべての点 e が等確率ならば，$x(e)$ は確率変数であることを証明せよ．またその標本空間を述べ，$P(x(e) = x')$ を計算せよ．

1.14 問題 1.4 において，N 個の電球を調べるときの，あらゆる可能な場合が等確率

であるとしよう．x 番目の電球を調べているときに s 番目の欠陥電球 $(s \leqslant r)$ が見つかる確率はどうか？ 求める確率が正となるための x の値の範囲を求めよ．

1.15 コインを続けて n 回投げる．この操作から生じる標本空間 R の 2^n 個の点に含まれる標本点を e とする．$x(e)$ を 表が出た回数とする．R の標本点にすべて等確率を定めると，$x(e)$ は確率変数となることを証明し，その標本空間を求めよ． また $P(x(e) = x')$ を計算せよ．

1.16 R は標本空間で，その要素 e は uv 平面において，頂点 $(0,0)$，$(0,1)$，$(1,0)$，$(1,1)$ を持つ正方形の内部の点とする． 座標 (u,v) を持つ任意の点 e に対し $x(e)=u+v$ とする．E を R 内の任意の 3 角形または 4 角形とし $P(E)=E$ の面積とすれば，$x(e)$ はすべての実数 x' に対し $P(x(e) \leqslant x')$ が計算できる確率変数であることを示せ．またその標本空間を述べ，$P(x(e) \leqslant x')$ に関する式を x' の関数として書け．F が $u/2 < v < u$ となる R の事象のとき，$\int_F x(e)\,dP$ の値を計算せよ． また $\int_R x(e)\,dP$ の値を計算せよ．

1.17 (続き) $y(e)=u/v$ とする． $y(e)$ はすべての実数 y' に関して $P(y(e) \leqslant y')$ を計算できる確率変数であることを示せ． また $P(y(e) \leqslant y')$ を計算せよ．

1.18 (続き) $(x(e), y(e))$ はあらゆる実数 x'，y' に対して $P(x(e) \leqslant x',\ y(e) \leqslant y')$ が計算できるような 2 次元の確率変数であることを示せ． また $P(x(e) \leqslant x',\ y(e) \leqslant y')$ を計算せよ．

1.19 (続き) E_n，$n=1,2,\cdots$ を $u>0,\ v>0,\ u+v \leqslant (1+n)/2n$ となる R 内の 3 角形とする．$\lim_{n\to\infty} E_n$ は何か？ また $\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n) = \dfrac{1}{8}$ であることを示せ．無限列 $E_1, E_2, \cdots$ は式 (1.9.3) を満足することを証明せよ．

1.20 **1.9.1** および **1.9.2** を証明せよ．

1.21 $(R, \mathscr{B}, P)$ を確率空間，$x(e)$ を $\mathscr{B}$ に関する非負の確率変数とする．$I_{\delta,\alpha}$ を区間 $(\alpha\delta, (\alpha+1)\delta)$，$\alpha=0,1,2,\cdots$，$\delta>0$，また $I_{\delta,\alpha}^{-1}$ を $x(e) \in I_{\delta,\alpha}$ となるような $\mathscr{B}$ の集合とする．$E \in \mathscr{B}$ に対し

$$A(\delta) = \sum_{\alpha=1}^{\infty} \alpha\delta P(E \cap I_{\delta,\alpha}^{-1})$$

とする．$A(\delta)$ が特定の δ の値に関して有限であるならば $\lim_{\delta\to 0} A(\delta)$ が存在して， それは E 上の $x(e)$ のルベーグ=スティルチェス積分であることを示せ．

1.22 $E_1, \cdots, E_n$ を任意の事象，$p_m(\alpha_1, \cdots, \alpha_r)$ を事象 $E_{\alpha_1}, \cdots, E_{\alpha_r}$ のうち少なくとも m 個が起る確率とする．このとき次のことを証明せよ．

$$(k+1-m) \sum\nolimits_{k+1} p_m(\alpha_1, \cdots, \alpha_{k+1}) \leqslant (n-k) \sum\nolimits_k p_m(\alpha_1, \cdots, \alpha_k).$$

ただし $k=1,\cdots,n-1,\ 1 \leqslant m \leqslant k$ で $\sum_i$，$i=k,\ k+1$ は $1,\cdots,n$ の中から i 個の整数を選び出す，$\binom{n}{i}$ 個の可能な選び方における和を示すものとする．(Chung (1941))．

第2章　分　布　関　数

2.1　は　じ　め　に

1.8 節で見たように，(1 次元) 確率変数 $x(e)$ は基本確率空間 $(R, \mathscr{B}, P)$ から，実直線 R_1 に関する確率空間 $(R_1, \mathscr{B}_1, P')$ を導いた． 同様にして， k 次元確率変数 $(x_1(e), \cdots, x_k(e))$ は基本確率空間 $(R,\ \mathscr{B},\ P)$ からユークリッド空間 R_k に関する確率空間 $(R_k,\ \mathscr{B}_k, P')$ を導いた．

R における $x(e)$ を扱う場合，通常，事象点 e を落して確率変数*) x と呼ぶ．x が $(R_1, \mathscr{B}_1, P')$ の (1 次元) 確率変数で，かつ E' が $\mathscr{B}_1$ の中の任意の集合であれば，あいまいな表現を避けるために， 事象 E' を $x \in E'$ で，$P'(E')$ を $P(x \in E')$ で表わす．特に E' が区間 $(a,b]$ であれば， $P(x \in E')$ は $P(a < x \leqslant b)$ であると解する． すなわち，いま E' が 1 点 x' だけであれば，$P(x=x')$ と表わす． これは与えられた試行において確率変数 x によって値 x' が表われたときの確率である． いま x が k 次元ベクトル確率変数 $(x_1, \cdots, x_k)$ で，x の次元を表わす必要があるときには，$\mathscr{B}_k$ の中の事象 E' を $(x_1, \cdots, x_k) \in E'$ で表わし， $P(x \in E')$ や $P'(E')$ ではなく $P((x_1, \cdots, x_k) \in E')$ を用いる． E' が k 次元の区間 $(a,b]_k$， すなわち $(a_1, \cdots, a_k\,;\, b_1, \cdots, b_k]$ ならば，$P((x_1, \cdots, x_k) \in E')$ は $P(a_i < x_i \leqslant b_i,\ i=1,\cdots,k)$ と書けることがわかる．

k 次元確率変数を論じるときには，“確率変数 $x_1, \cdots, x_k$” または “$x_1, \cdots, x_k$ を成分に持つ k 次元確率変数” と呼ぶこともある．

しばしば R_k の中の有界集合が問題になることがある．R_k の中の集合 E' が有界であ

---

*) 理想的には，確率変数を肉太の活字か特殊な記号にすると良い． しかし確率変数を表わす多くの異なった記号 (いくつかは古くさい) を含む確率変数の理論に関する応用の本では，実用的ではない． 実際には議論しているところが確率変数であるときには，この本ではどれが確率変数であるかすぐわかる．

るとは，E' がある有限 k 次元区間 $(a, b]_k$，ただし $a_1, \cdots, a_k ; b_1, \cdots, b_k$ はすべて有限である，に含まれる場合である．確率変数 $(x_1, \cdots, x_k)$ が有限区間 $(a, b]_k$ に対して，$P((x_1, \cdots, x_k) \in (a, b]_k) = 1$ の性質を持つとき，$(x_1, \cdots, x_k)$ は確率 1 で有界であるという．

2.2　1 変数の分布関数

x が 1 次元確率変数で，確率空間[*]が $(R_1, \mathscr{B}_1, P)$ であるとすれば，R_1 の中の x の標本空間上の確率分布は R_1 の中の各点で定義されている**分布関数** $F(x)$ で表現でき，特定の性質を持つことを示そう．

$\mathscr{B}_1$ に属する R_1 上の任意の区間 $(-\infty, x']$ に対して，$F(x')$ を次のように定義する．

(2.2.1)
$$F(x') = P(-\infty < x \leqslant x').$$

$F(x)$ は明らかに 1 価で，実数値をとり，R における x の非負の関数である．

$x'' > x'$ であれば，**1.4.2** から

(2.2.2)
$$\begin{aligned} F(x'') - F(x') &= P(-\infty < x \leqslant x'') - P(-\infty < x \leqslant x') \\ &= P(x' < x \leqslant x'') \geqslant 0. \end{aligned}$$

よって，$F(x)$ は x の非減少関数である．

区間 $(-\infty, \alpha]$ を E_α で表わすと，次のような集合の減少列を得る．

$$E_{-1} \supset E_{-2} \supset \cdots.$$

これから明らかに，$\lim_{\alpha \to \infty} E_{-\alpha} = \varnothing$ となる．そして **1.4.5** から

$$F(-\infty) = \lim_{\alpha \to \infty} F(-\alpha) = \lim_{\alpha \to \infty} P(E_{-\alpha}) = P(\lim_{\alpha \to \infty} E_{-\alpha}) = P(\varnothing) = 0$$

すなわち

(2.2.3)
$$F(-\infty) = 0.$$

同様に，増加集合列であれば $E_1 \subset E_2 \subset \cdots$ となり，これから $\lim_{\alpha \to \infty} E_\alpha = R_1$ となる．ゆえに **1.4.5** から

$$F(+\infty) = \lim_{\alpha \to \infty} F(\alpha) = \lim_{\alpha \to \infty} P(E_\alpha) = P(\lim_{\alpha \to \infty} E_\alpha) = P(R_1) = 1$$

すなわち

(2.2.4)
$$F(+\infty) = 1.$$

[*] 今後，ことわらないかぎり，E (E' ではない) は $\mathscr{B}_1$ の中の集合を，さらに一般的には $\mathscr{B}_k$ の中の集合を表わし，P にはダッシュをつけない．

ここで，$\lim_{\alpha \to \infty} x_\alpha = x'$ のような実数 $x_1, x_2, \cdots$ の減少列を考える．このとき，$E_{x_1}, E_{x_2}, \cdots$ は極限 $E_{x'}$ を持つ集合の減少列となる．ふたたび **1.4.5** を用いると

$$\lim_{\alpha \to \infty} F(x_\alpha) = \lim_{\alpha \to \infty} P(E_{x_\alpha}) = P(\lim_{\alpha \to \infty} E_{x_\alpha}) = P(E_{x'}) = F(x')$$

すなわち

(2.2.5)
$$F(x' + 0) = F(x').$$

いい換えると，ダッシュをとって，$F(x)$ は x の各値で右連続である．

読者は，(2.2.5) が半閉区間 $(-\infty, x']$ に含まれる確率として，$F(x')$ を定義していることに気づくであろう．$F(x')$ の定義に $P(-\infty < x < x')$ をとったとすると，$F(x' - 0) = F(x')$ という結果を得る．すなわち $F(x)$ は x の各値について左連続となる．$F(x')$ の定義は開区間 $(-\infty, x')$ ではなく，半閉区間 $(-\infty, x']$ に含まれる確率とし，(2.2.5) の関係を持っているものとした方が良い．

x が確率空間 $(R_1, \mathscr{B}_1, P)$ を持つ確率変数であるならば，R_1 の各点 x に対して，(2.2.1)（ダッシュを落した）で定義された関数 $F(x)$ が存在し，(2.2.2) から (2.2.5) までの性質を持っている．

逆に，関数 $F(x)$ が (2.2.1) の定義と (2.2.2) から (2.2.5) までの性質を持つように与えられたならば，確率空間 $(R_1, \mathscr{B}_1, P)$ が存在する．なぜならば，最初に考えた R_1 での集合族 $\mathscr{F}_0$ のかわりに，すべての実数 x に対する半開区間の族を使い，$(-\infty, x]$ に確率 $F(x)$ を割り当てると良い．この集合の初期族がボレル集合体 $\mathscr{B}(\mathscr{F}_0)$ を生成し，定義により $\mathscr{B}_1$ に等しくなる．$F(x)$ に関する (2.2.2) から (2.2.5) までの性質を利用して，$\mathscr{F}_0$ によって生成されるボレル集合体上に $F(x)$ から確率測度が一意に構成される．

まとめると

2.2.1　確率変数 x の確率空間 $(R_1, \mathscr{B}_1, P)$ は，R_1 の中の各点 x に対して，(2.2.1) によって定義された，1 価で実数値をとり，かつ非負の関数 $F(x)$ を一意に定める．この $F(x)$ は次の性質を持つ．

(2.2.6)
(a)　$F(x'') - F(x') \geqslant 0$　ただし　$x'' > x'$
(b)　$F(-\infty) = 0$
(c)　$F(+\infty) = 1$
(d)　$F(x + 0) = F(x)$.

逆に，関数 $F(x)$ がこれらの性質を持てば，$P(E_x) = F(x)$ なる確率空間

$(R_1, \mathscr{B}_1, P)$ が一意に定まる.

$F(x)$ を (1 次元) 確率変数 x の分布関数 (d.f.)[*] または累積分布関数 (c.d.f.)[**] と呼ぶ. 本書では後者を使うことにする. 全確率 1 が x 軸に分布していると考えれば, $F(x)$ は単に $(-\infty, x]$ 上にある確率の割合である.

このように (1 次元) 確率変数 x に関する確率を表わすには 2 通りの方法があることがわかった. 1 つは確率空間 $(R_1, \mathscr{B}_1, P)$ による方法で, もう 1 つは (2.2.6) に示した条件を満足する R_1 の各点に定義された c.d.f. による方法である. c.d.f. による表現は確率変数の解析に便利なので, 以後これを用いる. 基本標本空間 R が与えられていて, R の各点 e で確率変数 $x(e)$ が定義されているならば, 次のように定義された確率変数の c.d.f. $F(x)$ が存在する.

$$P(x(e) \leqslant x) = F(x).$$

逆に基本空間 R にまったく関係のない c.d.f. $F(x)$ が与えられたとする. そこでは基本標本空間 R として実数軸を考えれば, c.d.f. が $F(x)$ であるような確率変数 x を常に定義することができる. 標本点 e はそれゆえに実数で, 与えられた任意の実数 x' に対して, $x(x') = x'$ となるように確率変数 $x(e)$ を構成すれば $P(x(e) \leqslant x') = F(x')$ となる.

E が $\mathscr{B}_1$ の任意の集合であるとき, 確率 $P(x \in E)$ が存在して, かつ $F(x)$ から定めることができる. たとえば, 確率 $P(x' < x \leqslant x'')$ は (2.2.2) 式による $F(x)$ から定めることができる. 確率変数 x に関連した $F(x)$ の性質として (2.2.5) を導くとき用いた方法を使って, 次の関係式が求められる.

(2.2.7) $$P(x = x') = F(x') - F(x' - 0)$$

(2.2.8) $$P(x' < x < x'') = F(x'' - 0) - F(x')$$

(2.2.9) $$P(x' \leqslant x < x'') = F(x'' - 0) - F(x' - 0)$$

(2.2.10) $$P(x' \leqslant x \leqslant x'') = F(x'') - F(x' - 0).$$

x が有界な確率変数であれば, $a < b$ なる有限な a, b が存在する. ただし, a は $F(a) = 0$ となる最大の数で, b は $F(b) = 1$ となる最小の数で, $b - a$ を x の範囲という.

[*] d.f. は distribution function の略号. 今後はこの略号を用いる. (訳注)

[**] c.d.f. は cumulative distribution function の略号. (訳注)

2.3 1 変数の一般型

数理統計学上出て来る大部分の 1 次元確率変数には次の 2 通りの型がある. すなわち離散型および連続型である. 確率変数のこの 2 つの型に対して, 分布関数 $F(x)$ と確率測度 $P(E)$ は次のように, 2 つの関数の一方を用いて定義できる.

(a) 離散型

離散型の場合, c.d.f. $F(x)$ は階段関数である. すなわち, R_1 の中の点 $x^{(1)}, x^{(2)}, \cdots$ の有限または可付番無限個のところでのみ関数の値が変化する. そして, 有限な集積点を持たずに, そこでは $p(x^{(1)}), p(x^{(2)}), \cdots$ の大きさの値の跳びが起る. 跳び $p(x^{(\alpha)})$ は (2.2.7) で与えられ, 次のようになる.

(2.3.1) $$p(x^{(\alpha)}) = P(x = x^{(\alpha)}) = F(x^{(\alpha)}) - F(x^{(\alpha)} - 0).$$

R_1 の他のすべての点 x' では, $p(x') = P(x = x') = 0$ である. $F(x)$ は次のように跳びで表現することができる.

(2.3.2) $$F(x) = \sum_{x^{(\alpha)} \leqslant x} p(x^{(\alpha)}).$$

この和は $x^{(\alpha)} \leqslant x$ なるすべての α の値におよぶ. (2.3.2) で $x \longrightarrow +\infty$ とすると次のようになる.

(2.3.3) $$F(+\infty) = \sum_{\alpha} p(x^{(\alpha)}) = 1.$$

全確率 1 は点 $x^{(1)}, x^{(2)}, \cdots$ に分布している. これらの点を確率点または質点という. この点集合は (2.3.2) で与えられた c.d.f. を持つ確率変数 x の標本空間からなる.

離散型確率変数 x の場合, x の質点がどこにあるかは前後関係から明らかである. それゆえ, α を落して単に $p(x)$ と書いても, 関数 $p(x)$ と書いてもあいまいさはないであろう. 関数 $p(x)$ を x の確率関数 (p.f.)[*] と呼ぶ.

まとめると次のようになる.

2.3.1 離散型確率変数 x の c.d.f. $F(x)$ は p.f. $p(x)$ により一意に定まる. またこの逆も成り立つ.

[*] p.f. は probability function の略号. 今後はこの略号を用いる. (訳注)

離散型確率変数を含む問題では，一般に c.d.f. $F(x)$ よりも p.f. $p(x)$ で処理した方が便利である．

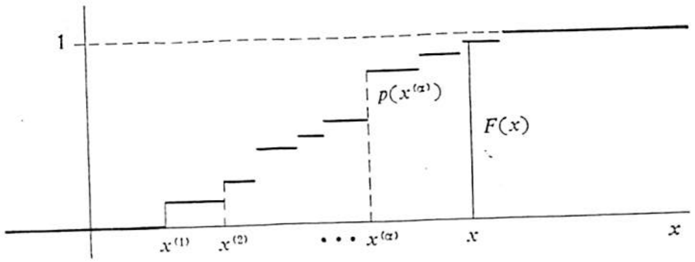

図 2.1　1 次元離散型確率変数 x の c.d.f. $F(x)$.

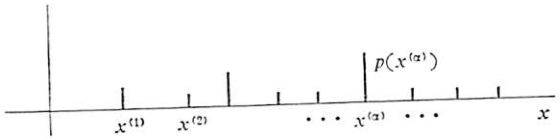

図 2.2　図 2.1 の c.d.f. $F(x)$ に対する p.f. $p(x)$.

1 次元離散型確率変数 x の c.d.f. $F(x)$ は図 2.1 に示したように，質点 $x^{(1)}, x^{(2)}, \cdots$ で大きさ $p(x^{(1)})$，$p(x^{(2)})$ の跳びのある階段関数で表わされる．

図 2.1 に示した c.d.f. を持つ確率変数の p.f. $p(x)$ は図 2.2 に示したように，質点 $x^{(1)}, x^{(2)}, \cdots$ のところでは垂直に $p(x^{(1)}), p(x^{(2)}), \cdots$ の長さだけあり，他の点ではみな 0 となるグラフに書ける．

$F(x)$ の $p(x^{(1)})$ という唯一の跳びが $x^{(1)}$ で起ったとすると，x は退化（1 次元）確率変数といい，c.d.f. を次のように書く．

(2.3.4)
$$\varepsilon(x-x^{(1)})=\begin{cases}1, & x\geqq x^{(1)}\\ 0, & x<x^{(1)}.\end{cases}$$

(2.3.1) で与えられた c.d.f. $F(x)$ は次に示すように ε 関数 (2.3.4) で表現できる．

(2.3.5)
$$F(x)=\sum_{\alpha}p(x^{(\alpha)})\cdot\varepsilon(x-x^{(\alpha)}).$$

例題　1 次元離散型確率変数の例は，初等的確率論には多数ある．たとえば，x を "理想的な" サイコロを投げたときの目の数を表わす確率変数とすれば，質点は $x^{(1)}=1, \cdots, x^{(6)}=6$ であり，確率は p.f. が $p(x)=\dfrac{1}{6}$，$x=1,\cdots,6$，x の他のすべての値で $p(x)=0$ となるように与えられる．52 枚のトランプを "良く切って" 13 枚配ったときのエースの枚数を x とすると，$x^{(1)}=0, x^{(2)}=1, \cdots, x^{(5)}=4$ となり，確率は p.f. が次を満たすように与えられる．

$$p(x)=\frac{\binom{4}{x}\binom{48}{13-x}}{\binom{52}{13}},\qquad x=0,1,2,3,4.$$

x の他のすべての値に対しては $p(x)=0$ である．

重要な離散型確率変数の例は第 6 章に示してある．

(b)　連　続　型

連続型確率変数に対しては，次のようなルベーグ可測関数 $f(x)\geqq 0$ がすべての $x'\in R_1$ に対して存在する．

(2.3.6)
$$F(x')=\int_{-\infty}^{x'}f(y)\,dy.$$

この場合，確率 0 の集合をのぞけば，dF/dx が存在し

(2.3.7)
$$\frac{dF}{dx}=f(x).$$

実際，関数 $f(x)$ は $F(x)$ が絶対連続のときのみ，(2.3.6) と (2.3.7) を満足するものが存在する．そしてこのような c.d.f. を持っている確率変数のことを絶対連続確率変数と呼ぶ．

$F(x)$ が (2.3.6) と (2.3.7) を満足しないで，単に連続な c.d.f. $F(x)$ を持つ確率変数 x を扱う場合がある．この確率変数の一般型は $F(x)$ が絶対連続の場合と間違わないようにしなくてはいけない．$F(x)$ が絶対連続でなくて，単に連続である場合はほとんどない．したがって通常，"絶対" という形容詞をはずして，x を連続型確率変数という．いかなる場合でも，どちらの連続型であるかは，前後関係から明らかになるであろう．たとえば第 8，11，13 章に出て来る順序統計量やノンパラメトリックの下での標本論では，$F(x)$ は単に連続である．次の式を見よう．

(2.3.8)
$$\frac{F(x'')-F(x')}{x''-x'}.$$

比 (2.3.8) は非負であり，(x',x'') での単位長さ当りの平均確率を表わしている．$x''\longrightarrow x'$ としたとき，極限が存在すれば，点 $x=x'$ における確率の密度として $f(x')\geqq 0$

が得られる．これを確率変数 x の確率密度関数（p.d.f.）[*] と呼び，次のようにまとめる．

2.3.2　連続型確率変数 x の c.d.f. $F(x)$ は (2.3.6) に従って p.d.f. $f(x)$ から一意に定まる．逆に，$f(x)$ は確率が 0 の集合をのぞいて (2.3.7) に従って $F(x)$ により定まる．

(2.3.7) が成り立てば，R_1 上の任意の（ボレル）集合 E に対して，次のようになる．

(2.3.9)
$$P(x \in E) = \int_E f(x)\,dx.$$

(2.3.9) からさらに進んで，$E = R_1$ であれば $\int_{R_1} f(x)\,dx = 1$ となる．常微分表示をすると次のように書ける．

(2.3.10)
$$P(x' < x < x' + dx) = f(x')\,dx.$$

もちろんここでは，dx は微分小を意味している．値 $f(x)\,dx$（ダッシュを落して）を x の確率素分（p.e. = probability element）という．$F(x)$ が単に連続であっても，便宜上，$P(x' < x < x' + dx)$ を $dF(x')$ で表わし，$x = x'$ における x の p.e. と呼ぶことにしよう．

区間 (0, 1) 上の任意の数 p に対して，c.d.f. $F(x)$ を持つ連続型確率変数 x の p 分位数（100 p 分位数）x_p は次を満たす最小数 $\underline{x}_p$ によって定義される．

(2.3.11)
$$F(\underline{x}_p) = p.$$

特に，$\underline{x}_{0.5}$ は x のメディアン[**]で，$\underline{x}_{0.25}$ と $\underline{x}_{0.75}$ はそれぞれ x の下 4 分位点，上 4 分位点である．離散型確率変数 x の場合，$F(x) = p$ を満たすような x が少なくとも 1 つあれば，p 分位数はこのような値の最小のものである．したがって，このようにして分位点は確率変数の質点のみにおいて定義される．

連続型確率変数 x を取り扱う問題では，c.d.f. $F(x)$ よりも p.d.f. $f(x)$ を使った方が便利である．

1 次元連続型確率変数の c.d.f. $F(x)$ とそれに対応する p.d.f. $f(x)$ は図 2.3 と図 2.4 に示してある．図 2.3 で $P(x' < x \leqslant x'')$ の値は 2 つの縦軸の高さ $F(x')$ と $F(x'')$ の差で表わされる．ところが図 2.4 では同じ確率が斜線の部分で表わされている．

---

[*] p.d.f. = probability density function で，今後この略号を用いる．（訳注）

[**] 中位数ともいう．（訳注）

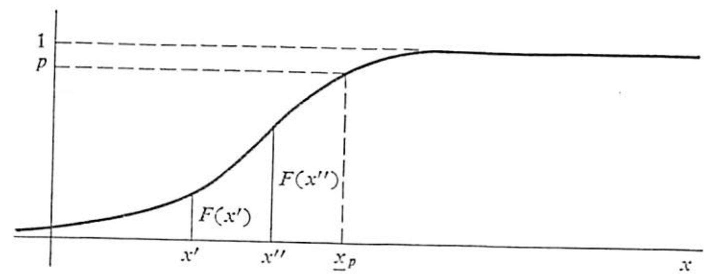

図 2.3　1 次元連続型確率変数 x の c.d.f. $F(x)$.

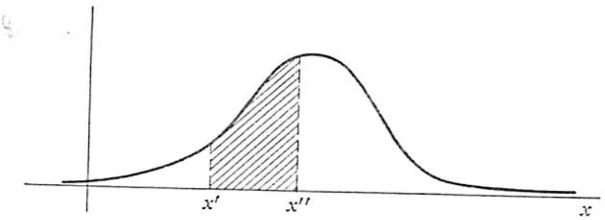

図 2.4　図 2.3 の c.d.f. $F(x)$ に対する p.d.f. $f(x)$.

例題　連続型確率分布の重要なものは第 7 章で取り上げる．ここでは次のような単純な例を考えよう．与えられた半径 r の円 C の中にある半径 δ の任意の円に，“ランダムな”点がはいる確率を δ^2/r^2 とする．円 C の中心から“ランダムな”点までの距離をとり，確率変数 x をこの距離と定義する．与えられた x' に対して $F(x')$ を $P(x \leqslant x')$ と定義すると，x（C の中に落ちる点）の c.d.f. は次のように与えられる．

$$F(x) = \begin{cases} 1 & x > r \\ \dfrac{x^2}{r^2} & 0 < x \leqslant r \\ 0 & x \leqslant 0. \end{cases}$$

そして x の p.d.f. は次のようになる．

$$f(x) = \begin{cases} \dfrac{2x}{r^2} & 0 < x \leqslant r \\ 0 & x \leqslant 0,\ x > r. \end{cases}$$

これで $F(x)$ の最も一般的な形は離散型 c.d.f. $F_1(x)$ と連続な（絶対連続である必要はない）c.d.f. $F_2(x)$ の凸結合であることが示せる．すなわち $F(x) = aF_1(x) + bF_2(x)$，ここで a，b は共に非負で $a + b = 1$ である．

2.4 2 変数の分布関数

(a) 一般的性質

(x_1, x_2) は確率空間 $(R_2, \mathscr{B}_2, P)$ を持つ 2 次元確率変数である．$E_{(x_1,x_2)}$ を R^2 の中の区間 $(-\infty, -\infty\,;\, x_1, x_2]$ とする．

(2.4.1) $F(x_1', x_2') = P(E_{(x_1',x_2')}) = P(-\infty < x_1 \leqq x_1', -\infty < x_2 \leqq x_2')$

とすると $F(x_1, x_2)$ は明らかに R_2 の中の (x_1, x_2) の 1 価で，実数値非負関数である．

$(x_1', x_2'\,;\, x_1'', x_2'']$ の型をした任意の区間 I_2 は

(2.4.2) $I_2 = (E_{(x_1'',x_2'')} - E_{(x_1',x_2'')}) - (E_{(x_1'',x_2')} - E_{(x_1',x_2')})$

であるから I_2 が $\mathscr{B}$ に属する．よって $(x_1, x_2) \in I_2$ の確率は次のようになる．

(2.4.3) $P((x_1, x_2) \in I_2) = F(x_1'', x_2'') - F(x_1', x_2'') - F(x_1'', x_1') + F(x_1', x_2')$.

(2.4.3) 式の右辺を $\Delta^2_{I_2}F(x_1, x_2)$ で表わし，I_2 上の $F(x_1, x_2)$ の 2 次の階差という．そして次のように表わす．

(2.4.4) $P((x_1, x_2) \in I_2) = \Delta^2_{I_2}F(x_1, x_2) \geqq 0.$

図 2.5 は $\Delta^2_{I_2}F(x_1, x_2)$ の各値を示している．$F(x_1', x_2'')$ と $F(x_1'', x_2')$ はそれぞれ縦線と横線を引いた無限領域（境界を含めて）に含まれる確率である．$F(x_1', x_2')$ は縦横両方の線の引かれた領域（上側と右側の境界を含めて）にはいる確率で，$\Delta^2_{I_2}F(x_1, x_2)$ は白い部分の矩形の領域（上側と右側の境界を含めて）にはいる確率である．

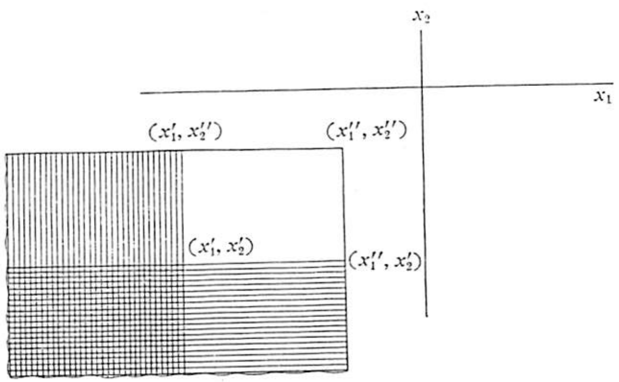

図 2.5 $\Delta^1_{I_2}F(x_1, x_2)$ に対する図.

さて，次の集合列を考えよう．

$$E_{(-\alpha, x_2')}, \qquad \alpha = 1, 2, \cdots.$$

これは $\lim_{\alpha\to\infty} E_{(-\alpha, x_2')} = \varnothing$ となる．それゆえ，**1.4.5** から，

$$\lim_{\alpha\to\infty} F(-\alpha, x_2') = \lim_{\alpha\to\infty} P(E_{(-\alpha, x_2')}) = P(\lim_{\alpha\to\infty} E_{(-\alpha, x_2')}) = P(\varnothing) = 0.$$

ここでダッシュを落して

(2.4.5) $F(-\infty, x_2) = 0.$

同様に

(2.4.5 a) $F(x_1, -\infty) = 0.$

さて次の集合列を考えよう．

$$E_{(\alpha,\alpha)}, \qquad \alpha = 1, 2, \cdots.$$

これは $\lim_{\alpha\to\infty} E_{(\alpha,\alpha)} = R_2$ となる増加集合列である．したがって **1.4.5** から

$$\lim_{\alpha\to\infty} F(\alpha, \alpha) = \lim_{\alpha\to\infty} P(E_{(\alpha,\alpha)}) = P(\lim_{\alpha\to\infty} E_{(\alpha,\alpha)}) = P(R_2) = 1$$

となり，これは

(2.4.6) $F(+\infty, +\infty) = 1.$

(2.2.5) を求めるのと同じ方法で，$F(x_1, x_2)$ が各変数に関して右連続であることがわかる．すなわち R_2 の中の各点 (x_1, x_2) で

(2.4.7) $F(x_1 + 0, x_2) = F(x_1, x_2 + 0) = F(x_1, x_2).$

(2.4.7) から $F(x_1 + 0, x_2 + 0) = F(x_1, x_2)$ が示せる．

したがって，(x_1, x_2) が確率空間 $(R_2, \mathscr{B}_2, P)$ を持つ確率変数であれば，R_2 の各点 (x_1, x_2) で (2.4.1) によって定義され，(2.4.4) から (2.4.7) までの性質を持った関数 $F(x_1, x_2)$ が一意に存在する．

逆に，1 変数の場合と同じように，(2.4.1) によって定義された関数 $F(x_1, x_2)$ が (2.4.4) から (2.4.7) までの性質を持てば，確率空間 $(R_2, \mathscr{B}_2, P)$ が一意に定まる．まとめると，**2.2.1** の 2 変数の場合が得られる．

2.4.1 2 次元確率変数 (x_1, x_2) の確率空間 $(R_2, \mathscr{B}_2, P)$ は R_2 における各点 (x_1, x_2) で，(2.4.1) で定義された 1 価で，実数値非負関数 $F(x_1, x_2)$ を一意に定める．しかも次の性質を持っている．

(2.4.8) (a) $\Delta^2_{I_2}F(x_1, x_2) \geqq 0$

(b) $F(-\infty, x_2) = F(x_1, -\infty) = 0$

(c) $F(+\infty, +\infty) = 1$

(d) $F(x_1 + 0, x_2) = F(x_1, x_2 + 0) = F(x_1, x_2)$.

逆に上の性質を持つ関数は $P(E_{(x_1, x_2)}) = F(x_1, x_2)$ となる確率空間 $(R_2, \mathscr{B}_2, P)$ を一意に定める．

$F(x_1, x_2)$ を 2 次元確率変数 (x_1, x_2) の分布関数または累積分布関数という．または 2 変数 x_1 と x_2 の c.d.f. であるという．すなわち $F(x_1, x_2)$ は 2 変数 c.d.f. といわれる．

1 次元確率変数の場合と同様に，2 次元確率変数の確率分布の表現に，確率空間と c.d.f. の 2 つの方式がある．本書では後者を用いる．

公式 (2.2.7) から (2.2.10) までの 2 次元確率変数の場合は読者自身で導き出し証明せよ．特に

(2.4.9) $$P(x_1 = x_1', x_2 = x_2') = F(x_1', x_2') - F(x_1' - 0, x_2') - F(x_1', x_2' - 0) + F(x_1' - 0, x_2' - 0).$$

(b) 周辺分布

次の集合を考えよう．

$$E_{(x_1', \alpha)}, \qquad \alpha = 1, 2, \cdots.$$

これは次のような増加集合列になっている．

$$\lim_{\alpha \to \infty} E_{(x_1', \alpha)} = E_{(x_1', +\infty)} = E_{x_1'}.$$

ここで $E_{x_1'}$ は R_2 の中の $x_1 \leqq x_1'$ なる事象である．ここで **1.4.5** を使うと次のようになる．

$$\lim_{\alpha \to \infty} F(x_1', \alpha) = \lim_{\alpha \to \infty} P(E_{(x_1', \alpha)}) = P(\lim_{\alpha \to \infty} E_{(x_1', \alpha)}) = P(E_{x_1'}).$$

すなわち

(2.4.10) $$F(x_1', +\infty) = P(E_{x_1'}) = P(x_1 \leqq x_1').$$

ダッシュを落して

(2.4.11) $$F(x_1, +\infty) = F_1(x_1)$$

とすると，$F_1(x_1)$ は 1 次元確率変数に対する c.d.f. の (2.2.6) (a) から (d) までの条件をすべて満足する．これは読者自身で証明せよ．事実，$F_1(x_1)$ は変数 (x_1, x_2) の成分 x_1 の c.d.f. である．そして，x_1 の周辺 c.d.f. または単に x_1 の c.d.f. という．

同様に

(2.4.12) $$F_2(x_2) = F(+\infty, x_2)$$

は x_2 の周辺 c.d.f. である．全確率 1 が，c.d.f. $F(x_1, x_2)$ に応じて (x_1, x_2) 平面に分布している．またこの確率を x_1 軸 $R_1^{(1)}$ 上に直角に射影すると，$F_1(x_1)$ は x_1 軸上の $(-\infty, x_1]$ にある確率である．もちろん同じことが $F_2(x_2)$ に対しても成り立つ．

2.2.1 から次のことがいえる．すなわち周辺 c.d.f. の $F_1(x_1)$ と $F_2(x_2)$ はそれぞれ確率空間 $(R_1^{(1)}, \mathscr{B}_1^{(1)}, P^{(1)})$，$(R_1^{(2)}, \mathscr{B}_1^{(2)}, P^{(2)})$ を定める．

(c) 統計的独立確率変数

(x_1, x_2) が確率空間 $(R_2, \mathscr{B}_2, P)$ を持つ 2 次元確率変数で，かつ成分 x_1 と x_2 がそれぞれ確率空間 $(R_1^{(1)}, \mathscr{B}_1^{(1)}, P^{(1)})$ と $(R_1^{(2)}, \mathscr{B}_1^{(2)}, P^{(2)})$ を持つ 1 次元確率変数であれば，x_1 と x_2 は統計的独立 (1.6 節を見よ) であるという．$E^{(1)}$ と $E^{(2)}$ がそれぞれ $R_1^{(1)}$ と $R_1^{(2)}$ における (ボレル) 集合であれば，$E^{(1)} \times E^{(2)}$ 型の $R_2 = R_1^{(1)} \times R_1^{(2)}$ におけるどんな集合 E に対しても，$P(E) = P^{(1)}(E^{(1)}) \cdot P^{(2)}(E^{(2)})$ となる．

"統計的" という言葉を落して，単に独立確率変数ともいう．

x_1 と x_2 の独立性を c.d.f. で表わすと次のようになる．

2.4.2 (x_1, x_2) が c.d.f. $F(x_1, x_2)$ を持つ確率変数のとき，x_1 と x_2 が独立であるための必要十分条件は次の通りである．

(2.4.13) $$F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2).$$

ここで $F_1(x_1)$，$F_2(x_2)$ はそれぞれ x_1，x_2 の周辺 c.d.f. である．

2.4.2 の証明．x_1 と x_2 の標本空間をそれぞれ $R_1^{(1)}$ と $R_1^{(2)}$ で表わす．すなわち

(2.4.14) $$R_2 = R_1^{(1)} \times R_1^{(2)}.$$

$E_{x_1'}$ と $E_{x_2'}$ をそれぞれ $R_1^{(1)}$，$R_1^{(2)}$ における事象 $x_1 \leqq x_1'$，$x_2 \leqq x_2'$ であるとすれば

(2.4.15) $$E_{(x_1', x_2')} = E_{x_1'} \times E_{x_2'}$$

となり，x_1，x_2 が独立であれば

(2.4.16) $$P(E_{(x_1', x_2')}) = P^{(1)}(E_{x_1'}) \cdot P^{(2)}(E_{x_2'})$$

となる．すなわち

(2.4.17)　$$F(x_1', x_2') = F_1(x_1')\cdot F_2(x_2').$$

逆に，$F(x_1, x_2)$，$F_1(x_1)$，$F_2(x_2)$ を R_2 における各事象点 (x_1', x_2') で (2.4.17) が成り立つような c.d.f. とすれば，すなわち $E_{(x_1', x_2')} = E_{x_1'} \times E_{x_2'}$ なる各集合に対しても (2.4.16) が成り立つとする．1.6 節から，確率空間 $(R_2, \mathscr{B}_2, P)$ は次のような性質を持ち，R_2 の中に一意に定まる．すなわち $E^{(1)} \times E^{(2)}$ 型の $\mathscr{B}_2$ の中の任意の集合 E に対して $P(E) = P^{(1)}(E^{(1)})\cdot P^{2}(E^{(2)})$ となる．ただし $(R_1^{(1)}, \mathscr{B}_1^{(1)}, P^{(1)})$，$(R_1^{(2)}, \mathscr{B}_1^{(2)}, P^{(2)})$ はそれぞれ $F_1(x_1)$，$F_2(x_2)$ で定められる確率空間である．これは x_1 と x_2 が独立であるということと同値である．

2.5　2変数の一般型

確率論と統計学で取り扱う 2 次元確率変数は 3 つのタイプ，すなわち，**離散型**，**連続型**，**混合型**，に分かれる．しかし混合型は他の 2 つに比べると出現の頻度はごくまれである．ここではこれらについてくわしく述べる．

(a)　離散型

離散型確率変数 (x_1, x_2) では，$F(x_1, x_2)$ は階段関数で，(2.4.9) の右辺は有限または可付番無限点の質点 $(x_1^{\alpha}, x_2^{\alpha})$，$\alpha = 1, 2, \cdots$ をのぞいては 0 である．そして R_2 において有限な集積点を持たない．このような質点では

(2.5.1)　$$P(x_1 = x_1^{(\alpha)}, x_2 = x_2^{(\alpha)}) = p(x_1^{(\alpha)}, x_2^{(\alpha)})$$

となり，さらに

(2.5.2)　$$\sum_{\alpha} p(x_1^{(\alpha)}, x_2^{(\alpha)}) = 1$$

となる．

逆に，$p(x_1^{(\alpha)}, x_2^{(\alpha)}) > 0$，$\alpha = 1, 2, \cdots$ でかつ R_2 の他のすべての点で $p(x_1, x_2) = 0$ となるような点列 $(x_1^{(\alpha)}, x_2^{(\alpha)})$ が与えられれば，$F(x_1, x_2)$ は次のように定義された階段関数となる．

(2.5.3)　$$F(x_1, x_2) = \sum p(x_1^{(\alpha)}, x_2^{(\alpha)}).$$

ただし，和は $x_i^{(\alpha)} \leqslant x_i$，$i = 1, 2$ に対する α のすべての値におよぶ．

質点が $(x_1^{(1)}, x_2^{(1)})$ だけのとき，確率変数 (x_1, x_2) は**退化**している．この場合 $p(x_1^{(1)}, x_2^{(1)}) = 1$ である．そして

(2.5.4)　$$F(x_1, x_2) = \varepsilon(x_1 - x_1^{(1)}, x_2 - x_2^{(1)}) = \begin{cases} 1, & x_i \geqslant x_i^{(1)}, \quad i = 1, 2 \\ 0, & \text{その他} \end{cases}$$

と書ける．もちろん (x_1, x_2) の成分の 1 つ，たとえば x_1 が (2.3.4) で定義した c.d.f. を持つ退化確率変数となることもある．ことわり書きがなければ，(x_1, x_2) のどの成分も退化しないと仮定する．

あいまいなところがなければ，α をとりのぞいて，$p(x_1, x_2)$ を (x_1, x_2) の確率関数 (p.f.) と呼ぶ．

まとめると次のようになる．

2.5.1　離散型確率変数 (x_1, x_2) の c.d.f. $F(x_1, x_2)$ は p.f. $p(x_1, x_2)$ により，一意に定まる．また逆も成り立つ．

離散型確率変数 (x_1, x_2) を取り扱うには，$F(x_1, x_2)$ よりも $p(x_1, x_2)$ の方が便利である．

一般に，E が R_2 における任意の集合であれば

$$P((x_1, x_2) \in E) = \sum_{(x_1^{(\alpha)}, x_2^{(\alpha)}) \in E} p(x_1^{(\alpha)}, x_2^{(\alpha)})$$

となる．特に x_1 の任意の値，すなわち x_1' に対する周辺 c.d.f. $F_1(x_1)$ は

(2.5.5)　$$F_1(x_1') = P(x_1 \leqslant x_1') = \sum_{(x_1^{(\alpha)}, x_2^{(\alpha)}) \in E_{x_1'}} p(x_1^{(\alpha)}, x_2^{(\alpha)})$$

で与えられる．ただし $E_{x_1'}$ は $x_1 \leqslant x_1'$ となる R_2 における集合である．$F_2(x_2)$ も上と同じように定義される．

周辺 c.d.f. $F_1(x_1)$ と $F_2(x_2)$ はおのおの周辺 p.f. $p_1(x_1)$，$p_2(x_2)$ を持つ離散型 1 次元 c.d.f. である．

次式の証明は容易なので読者に残す．

2.5.2　(x_1, x_2) が p.f. $p(x_1, x_2)$ を持つ離散型確率変数であれば，x_1 と x_2 が独立であるための必要十分条件は，$p(x_1, x_2) = p_1(x_1)\cdot p_2(x_2)$ である．

例題　2 次元離散型確率変数の例は初等的確率論に多数出て来る．たとえばトランプの 13 枚の持札において，x_1 をエースの枚数，x_2 をキングの枚数と定義すると，確率変数 (x_1, x_2) の質点 (x_1^α, x_2^α)，$\alpha = 1, 2, \cdots, 25$ は $(0,0)$，$(0,1)$，$(1,0), \cdots, (4,4)$ である．"完全に" 切ってあるという条件の下では，すなわち $\binom{52}{13}$ 組の手がみな等確率であれば，(x_1, x_2) の p.f. は

$$p(x_1, x_2) = \frac{\binom{4}{x_1}\binom{4}{x_2}\binom{44}{13 - x_1 - x_2}}{\binom{52}{13}}.$$

ここで，$(x_1, x_2) = (0,0), (0,1), \cdots, (4,4)$ 以外のすべての R_2 の点で $p(x_1, x_2) = 0$．x_1 の周辺 p.f. は

$$p_1(x_1) = \sum_{x_2=0}^{4} p(x_1, x_2) = \frac{\binom{4}{x_1}\binom{48}{13 - x_1}}{\binom{52}{13}}$$

で与えられる．これはもちろんエースのはいっている数を表わす確率変数 x_1 の p.f. である．$p_2(x_2)$ についても上と同じように表わせる．$p(x_1, x_2) \neq p_1(x_1)\cdot p_2(x_2)$ であるから，x_1 と x_2 は統計的に独立でない．

2 次元離散型確率変数の例は第 6 章で取り上げる．

(b) 連 続 型

2 次元連続型確率変数 (x_1, x_2) の場合には，次のようなルベーグ可測関数 $f(x_1, x_2) \geqslant 0$ が存在する．

(2.5.6)
$$F(x_1, x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(y_1, y_2)\, dy_1\, dy_2.$$

これはすべての $(x_1, x_2) \in R_2$ に対して成り立ち，$\dfrac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2}$ が存在し，

(2.5.7)
$$\frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} = f(x_1, x_2).$$

これは確率 0 の集合をのぞいた R_2 のすべての点で成り立つ．

(2.5.7) が意味を持つための $f(x_1, x_2)$ に関する条件は 1 次元確率変数の場合を 2 次元確率変数に拡張したのと同じである．その詳細は読者自身にまかせる．

さて次に 2 次元確率変数 (x_1, x_2) の確率密度の概念を考えてみよう．

次の比をとってみる．

(2.5.8)
$$\frac{\Delta_{I_2}^2 F(x_1, x_2)}{(x_1'' - x_1')(x_2'' - x_2')}, \qquad x_1'' > x_1', \qquad x_2'' > x_2'.$$

ただし分子は (2.4.4) に定義されている．この比は I_2 の中の単位区間あたりの平均確率を表わしている．$x_1'' \longrightarrow x_1'$ かつ $x_2'' \longrightarrow x_2'$ としたとき，比の極限が存在すれば，この極限は $f(x_1', x_2') \geqslant 0$ となる．すなわち (x_1', x_2') において非負の密度である．与えられた点で，$f(x_1, x_2)$ が (2.5.7) で示されるならば，(x_1, x_2) の確率密度がその点で存在するという．$f(x_1, x_2)$ を (x_1, x_2) の**確率密度関数** (p.d.f.)，また $f(x_1, x_2)\,dx_1\,dx_2$ を (x_1, x_2) の**確率素分** (p.e.) という．まとめると

2.5.3　連続型確率変数 (x_1, x_2) の c.d.f. $F(x_1, x_2)$ は (2.5.6) に与えられた p.d.f. $f(x_1, x_2)$ により一意に定まる．逆に $f(x_1, x_2)$ は確率 0 の R_2 の点集合をのぞいて，(2.5.7) で与えられた $F(x_1, x_2)$ により定まる．

E が R_2 の任意の（ボレル）集合であれば，$P(E)$ は次のルベーク積分で与えられる．

(2.5.9)
$$P((x_1, x_2) \in E) = \iint_E f(y_1, y_2)\, dy_1\, dy_2.$$

特に次のような周辺 c.d.f. を持つ．

(2.5.10)
$$F_1(x_1') = P(x_1 \leqslant x_1') = \int_{-\infty}^{x_1'}\int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2\, dy_1.$$

(2.5.10) の積分を重積分と考えれば，関数

(2.5.11)
$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, y_2)\, dy_2$$

は x_1 の周辺 p.d.f. と呼ばれる．x_2 の周辺 p.d.f. も同様に定義される．

次に示す **2.5.4** は 2 つの確率変数 x_1 と x_2 の統計的独立性の良い判定条件である．読者自身で確かめてみよ．

2.5.4　(x_1, x_2) が p.d.f. $f(x_1, x_2)$ を持つ連続型確率変数のとき，x_1, x_2 が統計的に独立であるための必要十分条件は

$$f(x_1, x_2) = f_1(x_1)\cdot f_2(x_2).$$

例題　2 次元連続型確率変数の簡単な例として，区間 $(0,1)$ から独立に 2 つの数を "取り出す" 試行を考える．この場合すべての数は "等確率" で取り出される．x_1，x_2 を，それぞれ取り出した小さい方の数，大きい方の数とすると，(2 次元) 確率変数となる．(x_1, x_2) の p.d.f. を次のように定義する．すなわち x_1x_2 平面上で $(0,0)$，$(1,1)$，

$(0,1)$ の 3 頂点を持つ 3 角形の中にあれば $f(x_1,x_2)=2$ とし，他の点では 0 とする．x_1 の周辺 p.d.f. は

$$f_1(x_1)=\int_{-\infty}^{\infty}f(x_1,y_2)\,dy_2=\int_{x_1}^{1}2\,dy_2=2(1-x_1)$$

となり，同様にして $f_2(x_2)=2x_2$ となる．

2 次元連続型確率変数の重要なものは第 7 章で取り上げる．

(c) 混合型

混合型確率変数 (x_1,x_2) では成分の 1 つが離散型で，もう 1 つが連続型である．さらに厳密にいうと，$F(x_1,x_2)$ は (x_1,x_2) の c.d.f. であり，かつ x_1 は離散型で x_2 は連続型である．そして $x_1^{(\alpha)}$, $\alpha=1,2,\cdots$ が x_1 の質点で，$p_1(x_1)$ が x_1 の p.f. であれば，1 次元条件つき p.d.f. $f(x_2|x_1^{(\alpha)})$, $\alpha=1,2,\cdots$ が存在し，

(2.5.12) $$f(x_2|x_1^{(\alpha)})=\frac{1}{p_1(x_1^{(\alpha)})}\cdot\frac{d}{dx_2}[F(x_1^{(\alpha)},x_2)-F(x_1^{(\alpha)}-0,x_2)]$$

で与えられる．そして

(2.5.13) $$F(x_1,x_2)=\sum_{x_1^{(\alpha)}\leqq x_1}p_1(x_1^{(\alpha)})\int_{-\infty}^{x_2}f(y|x_1^{(\alpha)})\,dy.$$

周辺 c.d.f. $F_1(x_1)$ と $F_2(x_2)$ は次のようになる．

(2.5.14) $$F_1(x_1)=\sum_{x_1^{(\alpha)}\leqq x_1}p_1(x_1^{(\alpha)})$$

(2.5.15) $$F_2(x_2)=\sum_{\alpha}p_1(x_1^{(\alpha)})\int_{-\infty}^{x_2}f(y|x_1^{(\alpha)})\,dy.$$

まとめると次のようになる．

2.5.5 x_1 は離散型で x_2 は連続型である混合型確率変数 (x_1,x_2) の c.d.f. $F(x_1,x_2)$ は，$[p(x_1^{(\alpha)}),f(x_2|x_1^{(\alpha)})]$, $\alpha=1,2,\cdots$ の組により一意に定まる．またこの逆も成り立つ．

いい換えると，x_1 が p.f. $p(x_1)$ を持つ離散型で，x_2 が条件つき p.d.f. $f(x_2|x_1)$ を持つ連続型である混合型確率変数 (x_1,x_2) は次のように表現できる．まず，確率 1 は $p_1(x_1^{(\alpha)})$, $\alpha=1,2,\cdots$ の大きさの部分に分割されており，かつその部分は x_1 軸上の質点 $x_1^{(\alpha)}$, $\alpha=1,2,\cdots$ に位置している．そして α 番目の直線上の任意の点 $(x_1^{(\alpha)},x_2)$ の

密度が $p_1(x_1^{(\alpha)})f(x_2|x_1^{(\alpha)})$ であるような方法で垂線 $x_1=x_1^{(\alpha)}$, $\alpha=1,2,\cdots$ にそって連続的に“ぬりつぶ”されている．

いま，$f(x_2|x_1^{(\alpha)})$ がすべての α に対して同じであれば

(2.5.16) $$f(x_2|x_1^{(\alpha)})=f_2(x_2)$$

となる．ただし $f_2(x_2)$ は x_2 の p.d.f. でかつ $F(x_1,x_2)=F_1(x_1)\cdot F_2(x_2)$，すなわち x_1 と x_2 は独立である．逆に x_1 と x_2 が独立であれば $F(x_1,x_2)=F_1(x_1)\cdot F_2(x_2)$ となり，かつ (2.5.12) から (2.5.16) を導き出すことができる．したがって，次が成り立つ．

2.5.6 (x_1,x_2) を x_1 が離散型で，x_2 が連続型の混合型確率変数とすれば，x_1 と x_2 が独立であるための必要十分条件はすべての α に対して $f(x_2|x_1^{(\alpha)})=f_2(x_2)$ である．ただし $f_2(x_2)$ は x_2 の周辺 p.d.f. である．

例題 前述のように，混合型確率変数は連続型，離散型に比べてかなりまれな場合であるが，次のような例で混合型確率変数を考えよう．サイコロを 1 つ振り，出た目の数を確率変数 x_1 で表わす．そして $x_1=x_1^{(\alpha)}$ ($x_1^{(1)}=1,\cdots,x_1^{(6)}=6$) ならば，$x_1^{(\alpha)}$ 個の数を区間 $(0,1)$ 上の“一様”分布から独立に取り出して，x_2 を $x_1^{(\alpha)}$ 個の出た目の数の中で 1 番大きなものとする．(x_1,x_2) は p.f. $p_1(x_1^{(\alpha)})=\dfrac{1}{6}$, $x_1^{(1)}=1,\cdots,x_1^{(6)}=6$ を持つ離散型確率変数 x_1 と次のような連続型確率変数 x_2 を持つ混合型確率変数となる．すなわち $0<x_2<1$ に対して

$$f(x_2|x_1^{(\alpha)})=\frac{d}{dx_2}(x_2)^{x_1^{(\alpha)}}=x_1^{(\alpha)}x_2^{x_1^{(\alpha)}-1}.$$

ただし $x_1^{(1)}=1,\cdots,x_1^{(6)}=6$ である．また $(0,1)$ 以外の x_2 のすべての値に対しては $f(x_2|x_1^{(1)})=0$ である．この例題の 1 つの分布は周辺 c.d.f. $F_2(x_2)$ でこれは次のように与えられる．

$$F_2(x_2)=\frac{1}{6}(x_2^1+x_2^2+\cdots+x_2^6).$$

これから p.d.f. は

$$f_2(x_2)=\frac{1}{6}(1+2x_2+\cdots+6x_2^5)$$

となることがわかる．混合型確率変数の例は第 8 章の問題 8.34 と 8.35 にもある．

2.6　k 変数の分布関数

(a) 一般的性質

2 次元確率変数に関する 2.4 と 2.5 節の結果を k 次元の場合に拡張することができる．その概略を示そう．多次元分布の詳細については von Neumann (1950) を参照されたい．

$(x_1, \cdots, x_k)$ が確率空間 $(R_k, \mathscr{B}_k, P)$ を持つ k 次元確率変数とする．$E_{(x_1', \cdots, x_k')}$ を R_k における集合 $(-\infty, \cdots, -\infty\,;\, x_1', \cdots, x_k']$ とすると，

(2.6.1)　$$F(x_1', \cdots, x_k') = P(E_{(x_1', \cdots, x_k')}) = P(-\infty < x_i \leqslant x_i',\ i = 1, \cdots, k).$$

$F(x_1, \cdots, x_k)$ は明らかに R_k における $(x_1, \cdots, x_k)$ の 1 価で，実数値非負関数である．

さて，$(x_1', \cdots, x_k'\,;\, x_1'', \cdots, x_k'']$ の形をした R_k の中の任意の区間 I_k は $\mathscr{B}_k$ に属するので

(2.6.2)　$$I_k = E_{(x_1'', \cdots, x_k'')} - [E_{(x_1', x_2'', \cdots, x_k'')} \cup \cdots \cup E_{(x_1'', \cdots, x_{k-1}'', x_k')}]$$

となる．

さらに $(x_1, \cdots, x_k) \in I_k$ となる確率は (2.4.3) を k 変数に拡張した，$F(x_1, \cdots, x_k)$ の項に見い出すことができる．これは

(2.6.3)　$$\begin{aligned} P((x_1, \cdots, x_k) \in I_k) = {} & F(x_1'', \cdots, x_k'') - [F(x_1', x_2'', \cdots, x_k'') + \cdots \\ & + F(x_1'', \cdots, x_{k-1}'', x_k')] + [F(x_1', x_2', x_3'', \cdots, x_k'') + \cdots \\ & + F(x_1'', \cdots, x_{k-2}'', x_{k-1}', x_k')] + \cdots + (-1)^k F(x_1', \cdots, x_k') \end{aligned}$$

で与えられる．この右辺を I_k 上の $F(x_1, \cdots, x_k)$ の k 次階差 $\Delta_{I_k}^k F(x_1, \cdots, x_k)$ で表わす．

(2.6.4)　$$P((x_1, \cdots, x_k) \in I_k) = \Delta_{I_k}^k F(x_1, \cdots, x_k) \geqslant 0.$$

(2.4.5) を求めたのと同じ方法で，

(2.6.5)　$$F(-\infty, x_2, \cdots, x_k) = \cdots = F(x_1, x_2, \cdots, x_{k-1}, -\infty) = 0.$$

また (2.4.6) を求めたのと同じ方法で

(2.6.6)　$$F(+\infty, \cdots, +\infty) = 1.$$

さらに (2.4.7) の k 次元の場合も容易に導ける．すなわち

(2.6.7)　$$F(x_1, \cdots, x_{i-1}, x_i + 0, x_{i+1}, \cdots, x_k) = F(x_1, \cdots, x_k), \qquad i = 1, \cdots, k.$$

これは x の任意の集合，すなわち $x_{i_1}, \cdots, x_{i_r}$ がそれぞれ $x_{i_1} + 0, \cdots, x_{i_r} + 0$ に置き

換わっても $F(x_1, \cdots, x_k)$ の値は変化しないことを表わしている．

それゆえ，$(x_1, \cdots, x_k)$ が確率空間 $(R_k, \mathscr{B}_k, P)$ を持つ k 次元確率変数であれば，(2.6.4) から (2.6.7) までの性質を持ち，(2.6.1) によって R_k の各点 $(x_1, \cdots, x_k)$ で定義される関数 $F(x_1, \cdots, x_k)$ が存在する．

逆に，$F(x_1, \cdots, x_k)$ が (2.6.1) で定義され，かつ (2.6.4) から (2.6.7) までの性質を持てば，確率空間 $(R_k, \mathscr{B}_k, P)$ を得る．

まとめると，**2.4.1** の k 次元への拡張は次のようになる．

2.6.1　k 次元確率変数 $(x_1, \cdots, x_k)$ の確率空間 $(R_k, \mathscr{B}_k, P)$ は R_k の中の各点で (2.6.1) で定義された，1 価で，実数値非負関数 $F(x_1, \cdots, x_k)$ を一意に定め，次の性質を持っている．

(2.6.8)
(a)　$\Delta_{I_k}^k F(x_1, \cdots, x_k) \geqslant 0$
(b)　$F(-\infty, x_2, \cdots, x_k) = \cdots = F(x_1, \cdots, x_{k-1}, -\infty) = 0$
(c)　$F(+\infty, \cdots, +\infty) = 1$
(d)　$F(x_1, \cdots, x_{i-1}, x_i + 0, x_{i+1}, \cdots, x_k) = F(x_1, \cdots, x_k)$,
　　$i = 1, \cdots, k.$

逆に，これらの性質を持ち，かつ $F(x_1, \cdots, x_k) = P(E_{(x_1, \cdots, x_k)})$ で定義された関数は確率空間 $(R_k, \mathscr{B}_k, P)$ を一意に定める．

関数 $F(x_1, \cdots, x_k)$ は k 次元確率変数 $(x_1, \cdots, x_k)$ の**分布関数** (d.f.) または累積分布関数 (c.d.f.) という．$F(x_1, \cdots, x_k)$ は，k 変数 c.d.f. ともいう．

$x_1 = x_1', \cdots, x_k = x_k'$ となる確率は (2.6.3) の右辺を $x_i'' \longrightarrow x_i'$, $i = 1, \cdots, k$ とした極限をとると得られることがわかる．すなわち

(2.6.9)　$$P(x_1 = x_1', \cdots, x_k = x_k') = \lim_{x_i'' \to x_i',\ \text{すべての}\ i} \Delta_{I_k}^k F(x_1, \cdots, x_k).$$

(b) 周 辺 分 布

x_1 の周辺 c.d.f. $F_1(x_1)$ は

(2.6.10)　$$F_1(x_1) = F(x_1, +\infty, \cdots, +\infty)$$

で定義される．他の周辺 c.d.f. $F_i(x_i)$, $i = 1, 2, \cdots, k$ も同様に定義される．さらに一般に $(x_1, \cdots, x_{k_1})$, $k_1 < k$ の周辺 c.d.f. は次のように定義される．

(2.6.11)　　$F_{1\cdots k_1}(x_1, \cdots, x_{k_1}) = F(x_1, \cdots, x_{k_1}, +\infty, \cdots, +\infty).$

もちろん，$(x_1, \cdots, x_k)$ の他の要素の任意の部分集合に対しても同じ定義が成り立つ．

(c) 2 つまたはそれ以上のベクトル確率変数の独立性

$(x_1, \cdots, x_{k_1})$，$(x_{k_1+1}, \cdots, x_k)$，(ただし $k = k_1 + k_2$) がそれぞれ確率空間 $(R_{k_1}, \mathscr{B}_{k_1}, P^{(1)})$，$(R_{k_2}, \mathscr{B}_{k_2}, P^{(2)})$ を持つベクトル確率変数であれば，2 つのベクトル確率変数は独立 (1.6 節を見よ) であるという．$E^{(1)}_{k_1}$，$E^{(2)}_{k_2}$ がおのおの $R^{(1)}_{k_1}$，$R^{(2)}_{k_2}$ のボレル集合で，$E^{(1)}_{k_1} \times E^{(2)}_{k_2}$ 型の $R_k = R^{(1)}_{k_1} \times R^{(2)}_{k_2}$ のどの集合 E_k に対しても，$P(E_k) = P^{(1)}(E_{k_1})\cdot P^{(2)}(E_{k_2})$ となる．

2 次元の場合については，独立性は次のような形で c.d.f. で表わすと便利である．読者自身で証明せよ．

2.6.2　$(x_1, \cdots, x_k)$ が c.d.f. $F(x_1, \cdots, x_k)$ を持つ確率変数であるとすれば，$(x_1, \cdots, x_{k_1})$ と $(x_{k_1+1}, \cdots, x_k)$ が独立であるための必要十分条件は次のようになる．

(2.6.12)　　$F(x_1, \cdots, x_k) = F_{1\cdots k_1}(x_1, \cdots, x_{k_1})\cdot F_{k_1+1\cdots k}(x_{k_1+1}, \cdots, x_k).$

ただし，右辺の 2 つの関数はそれぞれ $(x_1, \cdots, x_{k_1})$ と $(x_{k_1+1}, \cdots, x_k)$ の周辺 c.d.f. である．

この独立性の考え方は，$(x_1, \cdots, x_k)$ の成分が 3 つまたはそれ以上の互いに独立な部分集合になる場合に拡張できる．特に $x_1, \cdots, x_k$ が互いに独立であるための必要十分条件は

(2.6.13)　　$F(x_1, \cdots, x_k) = F_1(x_1)\cdots F_k(x_k).$

2.7 k 変数の一般型

2 次元の場合と同様に，k 次元確率変数にも 3 種類の一般型がある．すなわち離散型，連続型，混合型の確率変数である．

離散型の場合，$F(x_1, \cdots, x_k)$ は $(x_1^{(\alpha)}, \cdots, x_k^{(\alpha)})$，$\alpha = 1, 2, \cdots$ の有限または可付番無限点をのぞいた R_k のすべての点で，(2.6.9) の右辺が 0 となる階段関数である．そして

この点列は R_k において有限な集積点を持たない．この点では

(2.7.1)　　$P(x_1 = x_1^{(\alpha)}, \cdots, x_k = x_k^{(\alpha)}) = p(x_1^{(\alpha)}, \cdots, x_k^{(\alpha)}) > 0.$

かつ

(2.7.2)　　$\sum_{\alpha} p(x_1^{(\alpha)}, \cdots, x_k^{(\alpha)}) = 1.$

関数 $p(x_1, \cdots, x_k)$ は $(x_1^{(\alpha)}, \cdots, x_k^{(\alpha)})$，$\alpha = 1, 2, \cdots$ では (2.6.9) で与えられた値を，その他の R_k におけるすべての点では 0 の値をとり，(2.7.1) で示されている．すなわち $p(x_1, \cdots, x_k)$ は $(x_1, \cdots, x_k)$ の p.f. である．α の唯一の値，すなわち $\alpha = 1$ しか存在しない場合には，$p(x_1^{(1)}, \cdots, x_k^{(1)}) = 1$ であり，$(x_1, \cdots, x_k)$ は退化した確率変数である．(2.5.3) と **2.5.1** の拡張は簡単である．

読者は周辺 p.f. $p_i(x_i)$ が，または一般的に $p_{1\cdots k_1}(x_1, \cdots, x_{k_1})$ が，どのように定義されるか容易にわかるであろうし，**2.6.2** がどのようにして p.f. でいい換えられるかもわかるであろう．退化の定義はもちろん周辺分布にも拡張できる．特に重要なことは，$x_1, \cdots, x_k$ が互いに独立であるためには次の式が成り立つとき，かつそのときのみである．

(2.7.3)　　$p(x_1, \cdots, x_k) = p_1(x_1)\cdots p_k(x_k).$

k 次元連続型確率変数 $(x_1, \cdots, x_k)$ の場合には，非負のルベーグ可測関数 $f(x_1, \cdots, x_k)$ が存在し，次のようになる．すべての $(x_1, \cdots, x_k) \in R_k$ に対して

(2.7.4)　　$F(x_1, \cdots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f(y_1, \cdots, y_k)\, dy_1 \cdots dy_k.$

ここで

$$\frac{\partial^k F(x_1, \cdots, x_k)}{\partial x_1 \cdots \partial x_k}$$

が R_k で存在し，確率 0 の集合をのぞいた R_k におけるすべての点で

(2.7.5)　　$\dfrac{\partial^k F(x_1, \cdots, x_k)}{\partial x_1 \cdots \partial x_k} = f(x_1, \cdots, x_k)$

となる．

(2.7.5) が意味を持つための条件は 1 変数または 2 変数の場合を k 変数へ拡張したときの条件である．この拡張は読者にまかせる．

2.5.3 の k 変数への拡張も読者にまかせよう．

$(x_1, \cdots, x_k)$ の任意の成分 x_i の周辺 p.d.f. $f_i(x_i)$ または成分の任意の部分集合の p.d.f. $f_{1\cdots k_1}(x_1, \cdots, x_{k_1})$ は明らかに，対応する c.d.f. から定義される．**2.6.2** は連続型確率変数の場合には p.d.f. でいい換えることができる．さらに確率変数 $(x_1, \cdots, x_k)$ の成分 $x_1, \cdots, x_k$ が互いに独立であるのは次式が成り立つとき，かつそのときのみであ

る．

(2.7.6)　$$f(x_1,\cdots,x_k)=f_1(x_1)\cdots f_k(x_k).$$

$(x_1,\cdots,x_{k_1})$ が k_1 次元離散型確率変数で，かつ $(x_{k_1+1},\cdots,x_k)$, $k=k_1+k_2$ が k_2 次元連続型確率変数であれば，$(x_1,\cdots,x_k)$ は k 次元混合型確率変数である．このような確率変数は次式の条件つき p.d.f. $f(x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1})$ で定義される．

(2.7.7)　$$f(x_{k_1+1},\cdots,x_k|x_1^{(\alpha)},\cdots,x_{k_1}^{(\alpha)}) = \frac{1}{p_{1\cdots k_1}(x_1^{(\alpha)},\cdots,x_{k_1}^{(\alpha)})}\frac{\partial^{k_2}}{\partial x_{k_1+1}\cdots\partial x_k}[\Delta_{I_0(k_1)}^{k_1}F(x_1,\cdots,x_k)].$$

ただし，$p_{1\cdots k_1}(x_1,\cdots,x_{k_1})$ は $(x_1,\cdots,x_{k_1})$ の p.f. で，かつ

(2.7.8)　$$\Delta_{I_0(k_1)}^{k_1}F(x_1,\cdots,x_k)=\lim_{x_1'\to x_1,\cdots,x_{k_1}'\to x_{k_1}}[\Delta_{I_{k_1}}^{k_1}F(x_1,\cdots,x_k)]$$

である．$\Delta_{I_{k_1}}^{k_1}F(x_1,\cdots,x_k)$ は k_1 次元区間 $(x_1',\cdots,x_{k_1}';x_1,\cdots,x_{k_1}]$ 上で，$(x_1,\cdots,x_{k_1})$ に関する $F(x_1,\cdots,x_k)$ の k_1 次階差である．ただし，$x_{k_1+1},\cdots,x_k$ は固定してある．

2.5.5 と **2.5.6** は k 次元に簡単に拡張される．これは読者の演習とする．

2.8　確率変数の関数

(a)　1 次元確率変数の関数

2.2 節にあったように，(1 次元) 確率変数 x は確率空間 $(R_1,\mathscr{B}_1,P)$ を，あるいは同値的に c.d.f. $F(x)$ を持つ．今後，しばしば確率変数 x の可測関数*) $g(x)$ についての確率論を取り扱わねばならないだろう．すなわち，$g(x)$ は実数値をとり，1 価でかつ確率 0 の集合をのぞいた R_1 の各点で定義されていて，各実数 y に対して $g(x)\leqslant y$ となる R_1 の点集合は R_1 のボレル集合 $\mathscr{B}_1$ の族に属している．たとえば x が確率変数ならば，$\sin x$, e^x などの x の基本的な関数は，明らかに可測である．

すなわち $(R_1,\mathscr{B}_1,P)$ が確率変数 x の確率空間であること，および $g(x)$ の定義から，任意の実数 y に対して，$g(x)\leqslant y$ となる x の値の集合は $\mathscr{B}_1$ に含まれる．したがってこの集合の確率は $(R_1,\mathscr{B}_1,P)$ によって与えられる．

不等式 $g(x)\leqslant y$ の確率を，固定した y に対して，$H(y)$ で表わすと

*) このような関数は，ベール関数ともいう．

(2.8.1)　$$H(y)=P(g(x)\leqslant y)=P(x\in E_y).$$

ただし E_y は $g(x)\leqslant y$ となる (R_1 の中の) 点 x の集合を表わす．

$H(y)$ は c.d.f. としてのすべての性質を持つことが容易にわかる．したがって，

2.8.1　x が確率変数で，$g(x)$ が y によって示される確率変数であれば，(2.8.1) で示された $H(y)$ は y の c.d.f. である．

$g(x)$ が強い意味で単調で，かつある開区間 A のすべての x に対して連続な，0 でない微係数を持つとする．$y=g(x)$ とし，B を y 空間での A の像 (区間) とする．$x'\in A$ をとり $y'=g(x')$ とすると，ある $\Delta y>0$ に対して，等式 $g(x)=y'+\Delta y$ の唯一の解があり，それを $x=g^{-1}(y'+\Delta y)$ で表わす．

x の c.d.f. が $F(x)$ であれば

(2.8.2)　$$H(y'+\Delta y)-H(y')=\pm[F(g^{-1}(y'+\Delta y))-F(g^{-1}(y'))].$$

$g(x)$ が単調増加であればプラス符号で，単調減少の場合はマイナス符号をとる．

(2.8.2) の両辺を Δy で割ると次のように書ける．

(2.8.3)　$$\frac{H(y'+\Delta y)-H(y')}{\Delta y}=\left[\frac{F(g^{-1}(y'+\Delta y))-F(g^{-1}(y'))}{g^{-1}(y'+\Delta y)-g^{-1}(y')}\right] \times\left[\frac{\pm(g^{-1}(y'+\Delta y)-g^{-1}(y'))}{\Delta y}\right].$$

$g(x)$ が $x=x'$ で 0 でない微係数を持てば，$g^{-1}(y)$ は $y=y'$ で 0 でない微係数を持つ．すべての $x\in A$ に対して $F(x)$ が微係数 $f(x)$ を持つとすれば，$\Delta y\longrightarrow 0$ として，(2.8.3) の極限をとり，ダッシュを落して

$$\frac{dH(y)}{dy}=f(x)\cdot\left|\frac{dx}{dy}\right|$$

を得る．ただし右辺の x は $g^{-1}(y)$ で置き換えられる．プラス，マイナスのいずれかの符号を使うかわりに，簡単のために絶対値を用いる．さらに

(2.8.4)　$$\int_A f(x)\,dx=\int_B f(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|dy.$$

まとめると

2.8.2　x は p.d.f. $f(x)$ を持つ連続型確率変数で，$y=g(x)$ は強い意味で単調でかつある開区間 A で連続でしかも 0 でない導関数を持つ関数とする．また，B を y 空間での A の像とする．y が確率変数 $g(x)$ であれば，y は p.d.f.

$h(y)$ が B において存在する連続型確率変数であり,

(2.8.5) $$h(y) = f(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|$$

で与えられる．ただし $g^{-1}(y)$ は x に対する $g(x)=y$ の解である．さらに (2.8.4) も成り立つ．

"変換 $x=g^{-1}(y)$ は確率素分 $f(x)\,dx$ を確率素分 $h(y)\,dy$ に移す" またはより簡単に "$f(x)\,dx \longrightarrow h(y)\,dy$" ということもある．ただし $h(y)\,dy$ は (2.8.5) で与えられる．公式 (2.8.5) はもちろん，積分における被積分関数（ここでは確率密度関数）の積分変数を変える一般的な方法である．

例題　(2.8.5) で示した方法を実証するために，x は区間 $(0,a)$ の x に対して p.d.f. $f(x)=1/a$ で，他の x に対して p.d.f. $f(x)=0$ となるような連続型確率変数とする．確率変数 x^n の p.d.f. を求めたい．この場合，$g(x)=x^n$ かつ $g^{-1}(y)=y^{1/n}$. したがって，(2.8.5) を用いると $(0,a^n)$ の中の y に対して

$$h(y) = \frac{1}{na}\,y^{1/(n-1)}$$

となり，その他では 0 である．すなわち変換 $x=y^{1/n}$ に対して

$$\frac{1}{a}\cdot dx \longrightarrow \frac{1}{a}\cdot\left|\frac{dy^{1/n}}{dy}\right|\,dy = \frac{1}{na}\,y^{1/(n-1)}dy$$

を得る．

(b)　2 次元確率変数の関数

2 次元確率変数 (x_1,x_2) がある．x_1 と x_2 の関数 $g(x_1,x_2)$ が次の条件を満足すれば確率変数となる．すなわち $g(x_1,x_2)$ が実数値をとり，1 価でかつ確率 0 の集合をのぞいた R_2，すなわち x_1x_2 平面の各点で定義されている．そしてすべての実数 y に対して $g(x_1,x_2)\leqslant y$ となる R_2 の点集合が R_2 上のボレル集合族 $\mathscr{B}_2$ に属している場合である．$g(x_1,x_2)$ が確率変数であれば，それを y で表わし，$g(x_1,x_2)\leqslant y$ に対する x_1x_2 平面の点集合を E_y で表わす．

確率変数は確率空間 $(R_2,\mathscr{B}_2,P)$ または c.d.f. $F(x_1,x_2)$ により特徴づけられるから，$(x_1,x_2)\in E_y$ となる確率は $(R_2,\mathscr{B}_2,P)$（あるいは，もちろん任意の実数 y に対して $F(x_1,x_2)$）によって与えられる．$H(y)$ をこの確率とすれば，

(2.8.6) $$H(y) = P(g(x_1,x_2)\leqslant y) = P((x_1,x_2)\in E_y).$$

$H(y)$ は c.d.f.，すなわち確率変数 $y=g(x_1,x_2)$ の c.d.f. であることが証明できる．

確率変数 (x_1,x_2) のベクトル値関数 $(g_1(x_1,x_2),g_2(x_1,x_2))$ を考えよう．ただし $g_1(x_1,x_2)$, $g_2(x_1,x_2)$ は 1 価の実数値で，確率 0 となる集合をのぞいた R_2 のすべての点で定義されており，$g_i(x_1,x_2)\leqslant y_i$, $i=1,2$ となる x_1x_2 平面内の点集合は実数のすべての組 (y_1,y_2) に対して $\mathscr{B}_2$ に属している．与えられた y_1, y_2 に対し，この点集合を $E_{(y_1,y_2)}$ で表わせば，$(x_1,x_2)\in E_{(y_1,y_2)}$ となる確率は確率変数 (y_1,y_2) の確率空間 $(R_2,\mathscr{B}_2,P)$ によって表わされる．この確率を $H(y_1,y_2)$ とすれば

(2.8.7) $$H(y_1,y_2) = P(g_i(x_1,x_2)\leqslant y_i,\ i=1,2) = P((x_1,x_2)\in E_{(y_1,y_2)}).$$

$H(y_1,y_2)$ は 2 次元 c.d.f. の性質をすべて満足することが証明できる．よって

2.8.3　(x_1,x_2) が 2 次元確率変数で，(y_1,y_2) で確率変数 $(g_1(x_1,x_2),g_2(x_1,x_2))$ を表わせば，(2.8.7) で定義した $H(y_1,y_2)$ は (y_1,y_2) の c.d.f. である．

(y_1,y_2) の 1 つの成分 y_1 の c.d.f. は次で与えられる．

(2.8.8) $$H_1(y_1) = P(g_1(x_1,x_2)\leqslant y_1) = P((x_1,x_2)\in E_{y_1}).$$

ただし E_{y_1} は $g_1(x_1,x_2)\leqslant y_1$ となる x_1x_2 平面の点集合である．$H_1(y_1)$ について，実際には c.d.f. $H(y_1,y_2)$ において $y_2\longrightarrow+\infty$ として定義した y_1 の周辺 c.d.f. であるが，最初に $H(y_1,y_2)$ を決めないでも得られる．すなわち (2.8.8) を見れば，式は y_1 だけに関連があり，確率変数 $y_2=g_2(x_1,x_2)$ は無関係である．

2 次元確率変数 (x_1,x_2) の成分は，ともに 0 ではない実数の定数 c_1, c_2 が存在して確率変数 $c_1x_1+c_2x_2$ が退化する確率変数であるとき，**1 次従属**である．（通常，定数 c のどちらかが 0 の場合には，注意が必要である．この場合，成分 x_1 または x_2 のどちらか 1 つがそれ自身退化する．）c_1, c_2 はともに 0 でなく，$c_1x_1+c_2x_2$ が退化する確率変数であれば，x_1 と x_2 は**真に 1 次従属**であるという．その意味はある定数 c_3 に対して $P(c_1x_1+c_2x_2=c_3)=1$ であり，かつ直線 $c_1x_1+c_2x_2=c_3$ は両軸に対して平行にならない．(x_1,x_2) の成分の 1 つまたはそれ以上が退化する場合をのぞけば起りうる 1 次従属の型は 1 つで，かつそれは真の 1 次従属である．x_1, x_2 が 1 次従属でないときは，それらは **1 次独立**であるという．

特に次のように x_1-x_2 が退化する場合には

$$P(x_1-x_2=0)=1.$$

x_1, x_2 は同値な確率変数であるといわれる．$F(x_1, x_2)$ を 2 つの同値な確率変数 x_1, x_2 の c.d.f. とすれば，$F_1(x_1)$ と $F_2(x_2)$ は等しい．もちろん，この逆は必ずしも成り立たない．2 つのベクトル確率変数は，対応する要素が同値な確率変数のとき，同値であるという．

(c) 2 つの連続型確率変数の連続関数

(x_1, x_2) が連続型確率変数で，$y_i = g_i(x_1, x_2)$, $i = 1, 2$ が x_1x_2 平面から y_1y_2 平面への 1 対 1 変換である場合が，特に重要である．このとき，ある種の正則条件のもとでは，(x_1, x_2) の p.d.f. および $g_1(x_1, x_2)$, $g_2(x_1, x_2)$ の 1 次偏導関数で (y_1, y_2) の p.d.f. を明確に表現できる．正確にいうと次のようになる．

2.8.4　(x_1, x_2) を p.d.f. $f(x_1, x_2)$ を持つ連続型確率変数とし，$g_i(x_1, x_2)$, $i = 1, 2$ は x_1x_2 平面のある開領域 A で連続な 1 次偏導関数を持つ 1 価関数とする．$y_i = g_i(x_1, x_2)$, $i = 1, 2$ は A のすべての点で一意の逆数 $x_i = g_i^{-1}(x_1, x_2)$, $i = 1, 2$ を持ち，B を y_1y_2 平面での A の像とし，J を次の行列式で定義されたヤコビアンで A のすべての点で 0 でない値を持っているとする．

(2.8.9)
$$J = \left|\frac{\partial x_i}{\partial y_j}\right|, \qquad i, j = 1, 2.$$

このとき任意の点 $(y_1, y_2) \in B$ における (y_1, y_2) の p.d.f. $h(y_1, y_2)$ は次で与えられる．

(2.8.10)
$$h(y_1, y_2) = f(x_1, x_2) \cdot |J|.$$

ただし (2.8.10) 式の右辺で，(x_1, x_2) はそれぞれ

$$g_1^{-1}(y_1, y_2), \qquad g_2^{-1}(y_1, y_2)$$

で置き換えられる．

さらに

(2.8.11)
$$\int_A f(x_1, x_2)\, dx_1\, dx_2 = \int_B f(x_1, x_2)|J|\, dy_1\, dy_2.$$

確率密度関数に関する定理 **2.8.4** は，単に定理だけを述べておく．重積分の変数変換は微積分のテキストに出ている．この定理の証明は Widder (1947) を参照すると良い．

例題　(x_1, x_2) は次のような p.d.f. を持つ確率変数であるとする．

$$f(x_1, x_2) = \begin{cases} e^{-x_1 - x_2}, & x_1 \geqslant 0,\ x_2 \geqslant 0 \\ 0, & \text{その他}. \end{cases}$$

そして確率変数 $(x_1 + x_2, x_2/x_1)$ の p.d.f. を求めたい．ここで次のような変換をする．

$$y_1 = x_1 + x_2, \qquad y_2 = \frac{x_2}{x_1}.$$

逆変換は

$$x_1 = \frac{y_1}{1 + y_2}, \qquad x_2 = \frac{y_1 y_2}{1 + y_2}.$$

この変換は x_1x_2 平面の第 1 象限と y_1y_2 平面の第 2 象限の間の 1 対 1 写像である．第 1 象限のすべての点に関する変換のヤコビアンの絶対値は

$$\left|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right| = \frac{y_1}{(1 + y_2)^2}.$$

よって (y_1, y_2) の p.d.f. は

$$h(y_1, y_2) = \begin{cases} e^{-y_1}\dfrac{y_1}{(1 + y_2)^2}, & y_1 > 0,\ y_2 > 0 \\ 0, & \text{その他}. \end{cases}$$

さらに，y_1, y_2 は独立な確率変数であることがわかる (**2.5.4** を見よ)．

(d) k 次元確率変数の関数

上の結果を数個の確率変数のいくつかの関数の場合にそのまま拡張しよう．$(x_1, \cdots, x_k)$ が k 次元確率変数であるとき，$g(x_1, \cdots, x_k)$ が次の条件を満足すれば，確率変数となる．すなわち $g(x_1, \cdots, x_k)$ が実数値をとり，1 価関数で，確率 0 の集合をのぞいた R_k のすべての点で定義されていて，かつ $g(x_1, \cdots, x_k) \leqslant y$ となる R_k (x 平面) の点集合が y のすべての実数値に対して R_k のボレル集合 $\mathscr{B}_k$ に属する場合である．与えられた y に対して，E_y でこの点集合を示すとすれば，y の c.d.f. は

(2.8.12)
$$H(y) = P((x_1, \cdots, x_k) \in E_y).$$

1 次従属の定義を k 次元確率変数 $(x_1, \cdots, x_k)$ に拡張する．$c_1x_1 + \cdots + c_kx_k$ がすべてが 0 ではない定数 $c_1, \cdots, c_k$ に対して退化する確率変数であれば，$x_1, \cdots, x_k$ は **1 次従属**である．c_i のどれもが 0 でないとき，**真に 1 次従属**である．$x_1, \cdots, x_k$ が 1 次従属でないとき，**1 次独立**であるという．

$(y_1, \cdots, y_{k_1})$ で示される確率変数 $(x_1, \cdots, x_k)$ のベクトル値関数 $(g_1(x_1, \cdots, x_k), \cdots, g_{k_1}(x_1, \cdots, x_k))$ が次の条件を満足すれば k_1 次元確率変数である．すなわち，各成分 $g_1(x_1, \cdots, x_k), \cdots, g_{k_1}(x_1, \cdots, x_k)$ は実数値で，1 価関数で，確率 0 の集合をのぞいた

R_k のすべての点 $(x_1, \cdots, x_k)$ で定義されており，かつ任意の実数の組 $y_1, \cdots, y_{k_1}$ に対しての $g_i(x_1, \cdots, x_k) \leqslant y_i,\ i = 1, \cdots, k_1$ となる R_k の点集合 $E_{(y_1, \cdots, y_{k_1})}$ が $\mathscr{B}_k$ に属する場合である．よって事象 $E_{(y_1, \cdots, y_{k_1})}$ の確率は確率変数 $(x_1, \cdots, x_k)$ の確率空間 $(R_k, \mathscr{B}_k, P)$ または c.d.f. $F(x_1, \cdots, x_k)$ によって定まる．$(y_1, \cdots, y_{k_1})$ の c.d.f. $H(y_1, \cdots, y_{k_1})$ は次のように定義される．

(2.8.13)
$$H(y_1, \cdots, y_{k_1}) = P((x_1, \cdots, x_k) \in E_{(y_1, \cdots, y_{k_1})}).$$

さてここで，$k_1 = k$ でかつ $(x_1, \cdots, x_k)$ は p.d.f. $f(x_1, \cdots, x_k)$ を持つ連続型確率変数であるとしよう．x 平面のある開領域 A で，$y_i = g_i(x_1, \cdots, x_k),\ i = 1, \cdots, k$ が唯一の逆関数 $x_i = g_i^{-1}(y_1, \cdots, y_k),\ i = 1, \cdots, k$ を持つとする．ただし $g_i(x_1, \cdots, x_k)$ は A でヤコビアン $J = \left|\dfrac{\partial x_i}{\partial y_j}\right| \neq 0$ となるような連続な 1 次導関数を持つ．そして $(y_1, \cdots, y_k)$ の平面での A の像を B で表わす．そのとき，$(y_1, \cdots, y_k)$ は B の点 $(y_1, \cdots, y_k)$ で，次で与えられる p.d.f. を持つ連続な確率変数である．

(2.8.14)
$$h(y_1, \cdots, y_k) = f(x_1, \cdots, x_k) \cdot |J|$$

さらに

(2.8.15)
$$\int_A f(x_1, \cdots, x_k)\, dx_1 \cdots dx_k = \int_B f(x_1, \cdots, x_k)\, |J|\, dy_1 \cdots dy_k.$$

これらは (2.8.10) および (2.8.11) の k 次元への拡張である．$x_i = g_i^{-1}(y_1, \cdots, y_k),\ i = 1, \cdots, k$ で，x は (2.8.14) および (2.8.15) の右辺の y で表わされることがわかる．

2.9　条件つき分布関数

(a)　概　　説

1.9 節で取り上げた条件つき確率は標本空間 R における事象に対して定義された．基本的には確率変数によって定義された事象が重要だから，1.10 節の概念を事象が確率変数で定まる場合にあてはめると良い．x が確率空間 $(R_1, \mathscr{B}_1, P)$ または c.d.f. $F(x)$ を持つ確率変数であれば，G を $P(G) > 0$ となる $\mathscr{B}_1$ の任意の事象とし，$E_{x'}$ を $x \leqslant x'$ となる事象とする．

(2.9.1)
$$F(x'|G) = \frac{P(E_{x'} \cap G)}{P(G)}$$

とすると，$F(x|G)$ は c.d.f. に対する (2.2.6) の条件のすべてを満足し，$x \in G$ が与

えられたときの条件つき c.d.f. といわれる．逆に，**2.2.1** により，$F(x|G)$ は条件つき確率空間 $(R_1, \mathscr{B}_1, P)$ を一意に定める．

同様に，$(x_1, \cdots, x_k)$ を c.d.f. $F(x_1, \cdots, x_k)$ を持つ k 次元確率変数とする．G が R_k のボレル集合であれば，与えられた $(x_1, \cdots, x_k) \in G$ に対して，$(x_1, \cdots, x_k)$ の条件つき c.d.f. を次のように定義できる．

(2.9.2)
$$F(x_1, \cdots, x_k|G) = \frac{P(E_{(x_1, \cdots, x_k)} \cap G)}{P(G)}.$$

ただし $P(G) > 0$，$E_{(x_1, \cdots, x_k)}$ は k 次元区間 $(-\infty\,;\,x]_k$ である．

実際に (2.9.2) で重要なのは，事象 G が $(x_1, \cdots, x_k)$ の成分の 1 つまたはそれ以上を固定して得られる $(x_1, \cdots, x_k)$ 空間の切断からなる場合である．たとえば $(x_1, \cdots, x_k)$ が p.d.f. $f(x_1, \cdots, x_k)$ を持つ k 次元連続型確率変数の場合のように，(2.9.2) の右辺の分母，分子がともに 0 となるような状態になる．この問題は次のように処理する．

(b)　2 次元確率変数の条件つき分布関数

2 つの確率変数 x_1, x_2 があり，$x_1'' < x_1 \leqslant x_1'$ に対する R_2 の筒集合を I_1 とし，$P(I_1) > 0$ とする．

(2.9.2) から

(2.9.3)
$$F(x_1, x_2|I_1) = \frac{P(E_{(x_1, x_2)} \cap I_1)}{P(I_1)}.$$

$F(x_1, x_2)$ が (x_1, x_2) の c.d.f. であり，かつ $F_1(x_1)$ が x_1 の周辺 c.d.f. であれば，$F(x_1', x_2|I_1)$ は次のように書ける．

(2.9.4)
$$F(x_1', x_2|I_1) = \frac{F(x_1', x_2) - F(x_1'', x_2)}{F_1(x_1') - F_1(x_1'')}.$$

$F(x_1', x_2|I_1)$ は x_2 の関数として，c.d.f. である．$x_1'' \longrightarrow x_1'$ としたときの右辺の極限が存在すれば，それを $F(x_2|x_1')$ で表わす．すなわち

(2.9.5)
$$\lim_{x_1'' \to x_1'} F(x_1', x_2|I_1) = F(x_2|x_1').$$

各 x_2 に対して $F(x_2|x_1')$ が存在すれば，x_2 の関数として $F(x_2|x_1')$ は c.d.f. と考えられる．すなわち，(2.2.6) にあげた基本的な性質を持っている．これは $x_1 = x_1'$ が与えられたときの x_2 の条件つき c.d.f. という．一般にはダッシュを落して，$F(x_2|x_1)$ は連続型確率変数 $x_2|x_1$ の c.d.f. であるという．$F(x_2|x_1)$ は x_1 がパラメータの役割を果

たす c.d.f. である．x_1 は $F(x_2|x_1)$ の定義では確率変数ではないことを強調して，固定された変数ということもある．$x_2|x_1$ は x_1 の固定した値 x_1' に対する c.d.f. $F(x_2|x_1)$ が $x_1=x_1'$ となる x_1x_2 平面のすべての値に対して定義されている 1 次元確率変数とみなしても良い．つまり $F(x_2'|x_1')$ は直線 $x_1=x_1'$ にそって $x_2\leqslant x_2'$ となる確率（または確率密度）である．すなわち各直線にそって，全空間にしめる確率（または確率密度）の割合を表わしている．

これまでの説明はむしろ条件つき確率変数とその c.d.f. への接近であり，数理統計学上のすべての分布に適用できる．さらに一般的にはラドン＝ニコディムの定理 **1.10.1** を用いる．

2 次元の条件つき確率変数に 2 種類の型がある．それを次に示す．

タイプ A　この型は確率変数 (x_1,x_2) の成分 x_1 が離散型の場合である．すなわち $k=2$ に関する (2.9.2) で，x_1 を質点 $x_1^{(\alpha)}$，$\alpha=1,2,\cdots$ を持つ離散型確率変数とする．G を R_2 の集合で，$x_1=x_1^{(\beta)}$ となる (x_1,x_2) の標本空間とする．そうすれば

$$P(G)=P(x=x_1^{(\beta)})=P_1(x_1^{(\beta)})>0.$$

$k=2$ に関する (2.9.2) を用いれば，次式が得られる．

$$F(x_1^{(\beta)},x_2|x_1=x_1^{(\beta)})=\frac{P(E_{(x_1^{(\beta)},x_2)}\cap(x_1=x_1^{(\beta)}))}{P(x_1=x_1^{(\beta)})}. \tag{2.9.6}$$

これを $F(x_2|x_1^{(\beta)})$ で表わす．関数 $F(x_2|x_1^{(\beta)})$ は c.d.f. であり，**条件つき確率変数** $x_2|x_1^{(\beta)}$ の c.d.f. という．タイプ A の条件つき確率変数に関して，$k=2$ のときの (2.9.2) から直接 $F(x_2|x_1^{(\beta)})$ を定義し，G を $x_1=x_1^{(\beta)}$ なる R_2 の点集合に等しく定義できる．(2.9.5) の定義からも，I_1 の定義に $x_1'=x_1^{(\beta)}$ を選べば，$F(x_2|x_1^{(\beta)})$ に対して同じ結果を導き出せる．

$x_2|x_1^{(\beta)}$ は $x_1=x_1^{(\beta)}$ となる x_1x_2 平面の直線上に標本空間を持つ確率変数である．$x_2|x_1^{(\beta)}$ の c.d.f. は $F(x_2|x_1^{(\beta)})$ である．β を落して $F(x_2|x_1)$ で条件つき確率変数 $x_2|x_1$ の c.d.f. を表わす．

タイプ A の条件つき確率変数に 2 つの一般型がある．最も一般的な型は x_1，x_2 がともに離散型変数の場合である．この場合，

$$F(x_2|x_1^{(\beta)})=\sum_{x_2^{(\gamma)}\leqslant x_2}\frac{p(x_1^{(\beta)},x_2^{(\gamma)})}{p_1(x_1^{(\beta)})}. \tag{2.9.7}$$

ただし $p(x_1,x_2)$ は (x_1,x_2) の p.f. であり，$p_1(x_1)$ は x_1 の周辺 p.f. である．添字を落して，

$$p(x_2|x_1)=\frac{p(x_1,x_2)}{p_1(x_1)} \tag{2.9.8}$$

とする．この場合 $p(x_2|x_1)$ は条件つき確率変数 $x_2|x_1$ の p.f. である．与えられた x_1 座標を持つ少なくとも 2 つの相異なる質点がないならば，$x_2|x_1$ は退化する確率変数である．

(2.9.8) は次のように書き換えられる．

$$p(x_1,x_2)=p(x_2|x_1)\cdot p_1(x_1). \tag{2.9.9}$$

これは次に示すように (x_1,x_2) の p.f. を 2 段階で求める方法を示している．第 1 段階では $x_2|x_1$ の p.f. と x_1 の p.f. を求める．そして次にこの 2 つを掛け合わせる．問題となる x_2 の周辺 p.f. は通常 $p(x_1,x_2)$ から定まる．

例題　x_1 枚のエースがはいっているのが既知で，さらに x_2 枚のキングがはいるブリッジの手札の確率 $p(x_2|x_1)$ を求めたい．手札に x_1 枚のエースと x_2 枚のキングがはいっている確率は

$$p(x_1,x_2)=\frac{\binom{4}{x_1}\binom{4}{x_2}\binom{44}{13-x_1-x_2}}{\cdot\binom{52}{13}}$$

で与えられる．$p_1(x_1)=\sum_{x_2=0}^{4}p(x_1,x_2)$ は x_1 枚のエースのはいっている手札の確率であり，次のようになる．

$$p_1(x_1)=\frac{\binom{4}{x_1}\cdot\binom{48}{13-x_1}}{\binom{52}{13}}.$$

それゆえ (2.9.8) を用いて

$$p(x_2|x_1)=\frac{\binom{4}{x_2}\cdot\binom{44}{13-x_1-x_2}}{\binom{48}{13-x_1}}.$$

さて，(2.9.9) を実証する問題を考えよう．“ひずみのない” サイコロを投げて，出た目の数を確率変数 x_1 で表わし，$x_1=x_1'$ のとき x_1' 枚のコインを投げ，表が出た数を確率変数 x_2 で表わすことにする．問題は $p_2(x_2)$，すなわち全試行を通じて得られる表の数が x_2 である確率を求めることである．ここで

$$p(x_2|x_1) = \binom{x_1}{x_2}\left(\frac{1}{2}\right)^{x_1}, \qquad \text{かつ} \qquad p_1(x_1) = \frac{1}{6}.$$

よって
$$p(x_1, x_2) = \frac{1}{6}\binom{x_1}{x_2}\left(\frac{1}{2}\right)^{x_1}.$$

$p_2(x_2)$ を求めるために，$x_2, x_2+1, \cdots, 6$ の x_1 に関する $p(x_1, x_2)$ を合計すると，$x_2 = 0, 1, 2, 3, 4, 5, 6$ に対してそれぞれ次のようになる．

$$p_2(x_2) = \frac{63}{384}, \frac{120}{384}, \frac{99}{384}, \frac{64}{384}, \frac{29}{384}, \frac{8}{384}, \frac{1}{384}.$$

第2番目に，それほど一般的ではないが，タイプ A の条件つき確率変数で $F(x_2|x_1^{(\beta)})$ が x_2 に関して絶対連続である場合，密度関数 $f(x_2|x_1^{(\beta)})$ は次のようになる．

(2.9.10)
$$F(x_2|x_1^{(\beta)}) = \frac{F(x_1^{(\beta)}, x_2) - F(x_1^{(\beta)} - 0, x_2)}{p_1(x_1^{(\beta)})} = \int_{-\infty}^{x_2} f(y|x_1^{(\beta)})dy.$$

これは，$k=2$ としたときの (2.9.2) から，G を $x_1 = x_1^{(\beta)}$ となる R_2 の点集合としてとれば，すぐ得られる．また $x_1' = x_1^{(\beta)}$ と選べば，(2.9.5) からもまた得られる．

ここで $x_2|x_1^{(\beta)}$ は p.d.f. $f(x_2|x_1^{(\beta)})$ を持つ条件つき連続型確率変数である．この p.d.f. は $F(x_1, x_2)$ で表現されているとき，関数 $f(x_2|x_1^{(\beta)})$ を (2.5.12) で $\alpha = \beta$ として得られる．

タイプ B　この型は，x_1 は p.d.f. $f_1(x_1)$ を持つ連続型確率変数であり，2つの場合がある．最も一般的なのは，(x_1, x_2) が p.d.f. $f(x_1, x_2)$ を持つ連続型確率変数の場合で，(2.9.4) は次のようになる．

(2.9.11)
$$F(x_1', x_2|I_1) = \frac{\int_{x_1''}^{x_1'}\int_{-\infty}^{x_2} f(x_1, x_2)\, dx_2\, dx_1}{\int_{x_1''}^{x_1'} f_1(x_1)\, dx_1}.$$

$f_1(x_1)$ と $\int_{-\infty}^{x_2} f(x_1, x_2)\, dx_2$ は，x_1' で $f_1(x_1') > 0$ となる，x_1 の連続関数であれば，$x_1'' \longrightarrow x_1'$ として極限をとると次式が得られる．

(2.9.12)
$$F(x_2|x_1') = \frac{\int_{-\infty}^{x_2} f(x_1', x_2)\, dx_2}{f_1(x_1')}.$$

連続型の場合に関する $F(x_2|x_1')$ のこの表わし方は，離散型の場合の (2.9.7) と同じである．

(2.9.13)
$$f(x_2|x_1') = \frac{f(x_1', x_2)}{f_1(x_1')}$$

と置くと，$f(x_2|x_1')$ は条件つき確率変数 $x_2|x_1'$ の p.d.f. である．(2.9.13) は x_1' からダッシュを落して次のように書き換えられる．

(2.9.14)
$$f(x_1, x_2) = f(x_2|x_1)\cdot f_1(x_1).$$

これは，もちろん連続型確率変数に関する (2.9.9) と同じものである．

例題　x_1 を区間 $(0,1)$ から“ランダム”に取り出した数を表わす確率変数とする（ただし区間 $(0,1)$ 上のすべての数は“幾何学的な意味で同程度にありうる”ものとみなす）．そして x_2 を，x_1 の1つの値 x_1' があれば，区間 $(x_1', 1)$ からランダムに取り出した数を表わす確率変数とする．このとき x_2 の p.d.f. $f_2(x_2)$ を求めたい．

$$f_1(x_1) = \begin{cases} 1, & 0 < x_1 < 1 \\ 0, & \text{その他} \end{cases} \qquad f(x_2|x_1) = \begin{cases} \dfrac{1}{1-x_1}, & x_1 < x_2 < 1 \\ 0, & \text{その他} \end{cases}$$

ゆえに
$$f(x_1, x_2) = \begin{cases} \dfrac{1}{1-x_1}, & 0 < x_1 < x_2 < 1 \\ 0, & R_2 \text{ のその他}. \end{cases}$$

よって
$$f_2(x_2) = \int_0^{x_2} \frac{dx_1}{1-x_1} = -\log(1-x_2), \quad 0 < x_2 < 1.$$

$(0,1)$ 以外の x_2 の値に対して，$f_2(x_2) = 0$.

タイプ B の2番目の条件つき確率変数は，x_1 が連続型，x_2 が離散型である．$f_1(x_1)$ が $f_1(x_1') > 0$ となる x_1' で連続で，$F(x_1, x_2)$ と $F(x_1, x_2 - 0)$ が x_1' において x_1 の導関数を持つならば，(2.9.4) と (2.9.5) を応用すると次式が得られる．

(2.9.15)
$$F(x_2|x_1') = \sum_{x_2^{(\beta)} < x_2} p^*(x_2^{(\beta)}|x_1')$$

ただし

(2.9.16)
$$p^*(x_2^{(\beta)}|x_1') = \frac{f^*(x_1', x_2^{(\beta)})}{f_1(x_1')}$$

かつ

(2.9.17)
$$f^*(x_1, x_2) = \frac{\partial}{\partial x_1}[F(x_1, x_2) - F(x_1, x_2 - 0)].$$

$p^*(x_2|x_1)$ が確率（質点における）よりむしろ確率密度から構成されることをのぞけば，$p^*(x_2|x_1)$ が (2.9.8) の $p(x_2|x_1)$ と同じ構造をしていることがわかる．つまり確率 1 を $p_2(x_2^{(1)}), p_2(x_2^{(2)}), \cdots$ の部分に分割して，それぞれ密度関数（p.d.f. ではない）$f^*(x_1, x_2^{(1)}), f^*(x_1, x_2^{(2)}), \cdots$ に応じて，各直線 $x_2 = x_2^{(1)}, x_2 = x_2^{(2)}, \cdots$ にそって連続的にこれらの確率を“あてはめている”と考えて良かろう．

(c) k 次元確率変数への拡張

(2.9.5) で定義された条件つき c.d.f. はまったく同じ方法で多次元確率変数へと拡張できる. $(x_1, \cdots, x_k)$ を c.d.f. $F(x_1, \cdots, x_k)$ を持つ確率変数とする. $F_{1\cdots k_1}(x_1, \cdots, x_{k_1})$ を $k_1 < k = k_1 + k_2$ なる $(x_1, \cdots, x_{k_1})$ の周辺 c.d.f. とし, I_{k_1} を $x_i'' < x_i \leqslant x_i'$, $i = 1, \cdots, k_1$ となる R_k における筒集合とする. この筒集合の R_{k_1} 上, つまり $(x_1, \cdots, x_{k_1})$ 標本空間上への射影は k_1 次元区間 $(x_1'', \cdots, x_{k_1}''; x_1', \cdots, x_{k_1}']$ であり, これを I_{k_1} で表わす. $\Delta_{I_{k_1}}^{k_1} F_{1\cdots k_1}(x_1, \cdots, x_{k_1})$ を (2.6.3) と (2.6.4) で定義された I_{k_1} 上での $F_{1\cdots k_1}(x_1, \cdots, x_{k_1})$ の k 次階差とする. この k 次階差は $P(I_{k_1})$ に他ならない. $P(I_{k_1}) > 0$ と仮定する. さて

(2.9.18)
$$F(x_1, \cdots, x_k | I_{k_1}) = \frac{P(E_{(x_1, \cdots, x_k)} \cap I_{k_1})}{P(I_{k_1})}$$

と置くと, 特に

(2.9.19)
$$F(x_1', \cdots, x_{k_1}', x_{k_1+1}, \cdots, x_k | I_{k_1}) = \frac{\Delta_{I_{k_1}}^{k_1} F(x_1, \cdots, x_k)}{\Delta_{I_{k_1}}^{k_1} F_{1\cdots k_1}(x_1, \cdots, x_{k_1})}$$

となる. 右辺の分子は $(x_{k_1+1}, \cdots, x_k)$ の固定した値に対する $F(x_1, \cdots, x_k)$ の $(x_1, \cdots, x_{k_1})$ に関する k_1 次階差である.

$x_1'' \longrightarrow x_1', \cdots, x_{k_1}'' \longrightarrow x_{k_1}'$ としたとき (2.9.19) 式の右辺の極限が存在すれば,

(2.9.20)
$$\begin{aligned}\lim_{x_i'' \to x_i', i=1,2,\ldots,k_1} & F(x_1', \cdots, x_{k_1}', x_{k_1+1}, \cdots, x_k | I_{k_1}) \\ &= F(x_{k_1+1}, \cdots, x_k | x_1', \cdots, x_{k_1}')\end{aligned}$$

となり, ダッシュを落して, $F(x_{k_1+1}, \cdots, x_k | x_1, \cdots, x_{k_1})$ を条件つき確率変数 $(x_{k_1+1}, \cdots, x_k | x_1, \cdots, x_{k_1})$ の c.d.f. という. もしこれが存在すれば $F(x_{k_1+1}, \cdots, x_k | x_1, \cdots, x_{k_1})$ は k_2 次元 c.d.f. の性質 (2.6.8) のすべてを持つことが明らかにできる. (2.9.20) で極限が存在しない場合には, 1.10 節で説明したような一般的な接近方法が必要である.

離散型確率変数 $(x_1, \cdots, x_k)$ の場合, $x_1 = x_1', \cdots, x_{k_1} = x_{k_1}'$ における R_k の部分として, (2.9.2) の G を $P(G) > 0$ となるように選べば, 条件つき確率変数 $(x_{k_1+1}, \cdots, x_k | x_1', \cdots, x_{k_1}')$ が存在し, 次式で示す p.f. を持つ. これは (2.9.8) の拡張である.

(2.9.21)
$$p(x_{k_1+1}, \cdots, x_k | x_1', \cdots, x_{k_1}') = \frac{p(x_1', \cdots, x_{k_1}', x_{k_1+1}, \cdots, x_k)}{p_{1\cdots k_1}(x_1', \cdots, x_{k_1}')}.$$

ダッシュを落した $p(x_1, \cdots, x_k)$ は $(x_1, \cdots, x_k)$ の p.f. であり, $p_{1\cdots k_1}(x_1, \cdots, x_{k_1})$ は $(x_1, \cdots, x_{k_1})$ の周辺 p.f. である. ここで述べた条件の下では条件つき確率変数 $(x_{k_1+1}, \cdots, x_k | x_1', \cdots, x_{k_1}')$ [(2.9.21) 式の p.d.f. を持つ] の c.d.f. は (2.8.20) で与えられることを示せる.

$(x_1, \cdots, x_k)$ が離散型であれば, $(x_1, \cdots, x_k)$ における質点での p.f. は次のような積の形になることが証明できる.

(2.9.22)
$$p(x_1, \cdots, x_k) = p(x_k | x_1, \cdots, x_{k-1}) \cdot p_{1\cdots k-1}(x_{k-1} | x_1, \cdots, x_{k-2}) \cdots p_1(x_1).$$

$(x_1, \cdots, x_k)$ が連続型確率変数の場合, (2.9.21) に関して, 次のような p.d.f. となる.

(2.9.23)
$$f(x_{k_1+1}, \cdots, x_k | x_1, \cdots, x_{k_1}) = \frac{f(x_1, \cdots, x_k)}{f_{1\cdots k_1}(x_1, \cdots, x_{k_1})}.$$

ここで右辺の 2 つの関数は 2.7 節で定義した p.d.f. である. この p.d.f. は (2.9.12) を求めるのに用いた方法と仮定をそのまま拡張して (2.9.20) から求めることもできる.

連続型の場合の (2.9.22) 式に関しては次に示す p.d.f. の積の形で得られる.

(2.9.24)
$$f(x_1, \cdots, x_k) = f(x_k | x_1, \cdots, x_{k-1}) \cdot f_{1\cdots k-1}(x_{k-1} | x_1, \cdots, x_{k-2}) \cdots f_1(x_1).$$

ただし条件つき確率変数すべてに対して p.d.f. が存在するものとする.

(d) 独立な場合の条件つき分布関数

まず確率変数 (x_1, x_2) を考えよう. G を $P(G) > 0$ なる x_2 軸に平行な R_2 におけるボレル筒集合とする. G の点 (x_1, x_2) を $x_1 \in G$ と表わしても問題はない. $E_{(x_1', x_2')}$ は直積 $E_{x_1'} \times E_{x_2'}$ である. ただし $E_{x_1'}$ は $x_1 \leqslant x_1'$ に対する $R_1^{(1)}$ での集合であり, $E_{x_2'}$ は $x_2 \leqslant x_2'$ となる $R_1^{(2)}$ における集合である. すると

$$E_{(x_1', x_2')} \cap G = (E_{x_1'} \cap G) \times E_{x_2'}.$$

これと (2.9.2) から

(2.9.25)
$$F(x_1', x_2' | G) = \frac{P((E_{x_1'} \cap G) \times E_{x_2'})}{P(G)}.$$

x_1 と x_2 が独立であれば, (2.9.25) は次のようになる.

(2.9.26)
$$F(x_1', x_2' | G) = \frac{P(E_{x_1'} \cap G)}{P(G)} \cdot P(E_{x_2'}).$$

ゆえに次のような結論を得る (ダッシュを落して).

2.9.1　x_1 と x_2 が独立で, (x_1, x_2) が c.d.f. $F(x_1, x_2)$ を持つ確率変数であり,

かつ G が $P(G)>0$ である x_2 軸に平行な R_2 における任意の（ボレル）筒集合であれば

(2.9.27)　$$F(x_1,x_2|G)=F_1(x_1|G)\cdot F_2(x_2).$$

これは，G を x_2 軸に平行な筒集合として，$(x_1,x_2)\in G$ が与えられたとき，x_2 が任意の特定の集合に属しているという条件つき確率が G に関係がないという意味である．特に，G が (2.9.3) で使った $x_1\in I_1$ に対する集合であるとすれば，x_1 と x_2 が独立なら，**2.4.2** から (2.9.4) は次のようになる．

(2.9.28)　$$F(x_1',x_2|I_1)=F_2(x_2).$$

この場合，確率が 0 となる x_2 の集合をのぞけば，$\lim\limits_{x_1''\to x_1'} F(x_1',x_2|I_1)$ は存在し，$F_2(x_2)$ に等しくなる．ゆえに次の **2.9.1** の系を得る．

2.9.1 a　x_1 と x_2 が独立ならば，確率 0 となる集合をのぞいて

(2.9.29)　$$F(x_2|x_1)=F_2(x_2).$$

x_1 と x_2 が独立な場合，(2.9.8) は次のように変形できる．

(2.9.29 *a*)　$$p(x_2|x_1)=p_2(x_2).$$

そして (2.9.13) は

(2.9.29 *b*)　$$f(x_2|x_1)=f_2(x_2).$$

さらに一般化すると

2.9.2　$(x_1,\cdots,x_k)$ が c.d.f. $F(x_1,\cdots,x_k)$ を持つ確率変数であり，$(x_1,\cdots,x_{k_1})$ と $(x_{k_1+1},\cdots,x_k)$ が独立な確率変数で，G が $P(G)>0$ となるように，$x_1,\cdots,x_{k_1}$ の限定された値で定義された R_k の任意のボレル筒集合であれば，次のようになる．

(2.9.30)　$$F(x_1,\cdots,x_k|G)=F_{1\cdots k_1}(x_1,\cdots,x_{k_1}|G)\cdot F_{k_1+1,\ldots,k}(x_{k_1+1},\cdots,x_k).$$

G が (2.9.19) で使った $(x_1,\cdots,x_{k_1})\in I_{k_1}$ となる区間であり，さらに $(x_1,\cdots,x_{k_1})$ と $(x_{k_1+1},\cdots,x_k)$ が独立であるならば，$F(x_1,\cdots,x_k|I_{k_1})=F_{1,\ldots,k_1}(x_1,\cdots,x_{k_1}|I_{k_1})\cdot F_{k_1+1,\ldots,k}(x_{k_1+1},\cdots,x_k)$ となり，(2.9.19) は次のようになる．

(2.9.31)　$$F(x_1',\cdots,x_{k_1}',x_{k_1+1},\cdots,x_k|I_{k_1})=F_{k_1+1,\ldots,k}(x_{k_1+1},\cdots,x_k).$$

$x_1''\longrightarrow x_1',\cdots,x_{k_1}''\longrightarrow x_{k_1}'$ とすると，ほとんどいたるところで極限が存在し，**2.9.2** の次の系を得る．

2.9.2 a　$(x_1,\cdots,x_k)$ において $(x_1,\cdots,x_{k_1})$ と $(x_{k_1+1},\cdots,x_k)$ が独立で，c.d.f. $F(x_1,\cdots,x_k)$ を持つ確率変数であれば，確率 0 となる集合をのぞいて

(2.9.32)　$$F(x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1})=F_{k_1+1,\ldots,k}(x_{k_1+1},\cdots,x_k).$$

$(x_1,\cdots,x_{k_1})$ と $(x_{k_1+1},\cdots,x_k)$ が独立であれば，(2.9.32) 式から (2.9.21) と (2.9.23) は次のようになる．

(2.9.32 *a*)　$$p(x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1})=p_{k_1+1,\ldots,k}(x_{k_1+1},\cdots,x_k).$$

そして

(2.9.32 *b*)　$$f(x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1})=f_{k_1+1,\ldots,k}(x_{k_1+1},\cdots,x_k).$$

2.10　有限確率過程

k 次元確率変数 $(x_1,\cdots,x_k)$ は有限確率過程として取り扱われることもある．特に $(x_1,\cdots,x_k)$ の成分を，k 次元 c.d.f. が決まっている物理的操作の結果の測定に対応するように，応用した場合である．

例題　たとえば，x_1 を区間 (0, 1) から“ランダム”に取り出した数を表わす確率変数とする．すべての数は“等確率である”ものとする．x_2 を $(x_1, 1)$ から“ランダム”に取り出した数とし，k 個の数まで続ける．すると k 次元確率変数 $(x_1,\cdots,x_k)$ が得られ，この p.d.f. $f(x_1,\cdots,x_k)$ に (2.9.24) を適用すると次のようになる．すなわち $0<x_1<\cdots<x_k<1$ に対しては

$$f(x_1,\cdots,x_k)=\frac{1}{(1-k_{k-1})}\cdot\frac{1}{(1-x_{k-2})}\cdot\cdots\cdot\frac{1}{(1-x_1)}$$

となり，その他の場合には $f(x_1,\cdots,x_k)=0$ となる．$(x_1,\cdots,x_k)$ は上に示した方法で区間 (0, 1) を連続的に“切断”した結果を表わす（有限）確率過程となる．

読者は本書に有限確率過程の例が多数あることに気がつかれたと思う．無限確率過程，すなわち無限に多くの成分を持つ多次元確率変数もまたそこここに出て来る．これは第 4 章で定義し解説する．

問　　題

2.1　$F(x)$ が確率変数 x の c.d.f であるとき，適当な半開区間の列と **1.4.5** を考えて，(2.2.7)，(2.2.8)，(2.2.9)，(2.2.10) の各式を証明せよ．

2.2　適当な 2 次元半開区間の列を考えて，(2.4.9) を証明せよ．

2.3　$F_1(x)$ と $F_2(x)$ をそれぞれ離散型確率変数と連続型確率変数の c.d.f. とする．a，b が非負の数で和が 1 であれば，

$$aF_1(x)+bF_2(x)$$

は c.d.f. の性質をすべて満足することを示せ．

2.4　2 次元混合型確率変数 (x_1, x_2) は次のような性質を持っている．すなわち x_1 は p.f. $p(x_1)=\left(\frac{1}{2}\right)^{x_1}$，$x_1=1, 2, \cdots$ を持つ離散型確率変数で，x_2 は条件つき確率変数 $x_2|x_1$ が区間 $(0, 1)$ で p.d.f. $x_1(1-x_2)^{x_1-1}$ を持つ連続型確率変数である．条件のつかない確率変数 x_2 の p.d.f. は $(0, 1)$ 上で $2(1+x_2)^{-2}$ であることを示せ．

2.5　離散型確率変数 x_1 が $0<p<1$，$q=1-p$ のとき，p.f. qp^{x_1-1}，$x_1=1, 2, \cdots$ を持っている．そして x_2 は $x_2|x_1$ の p.d.f. が $(0, 1)$ 上で $x_1x_2^{x_1-1}$ で，その他の場合は 0 であるような連続型確率変数である．x_2 の条件のつかない p.d.f.，$x_1|x_2$ の p.f.，x_1 の c.d.f.，x_2 の c.d.f. を求めよ．

2.6　x_1x_2 平面の第 1 象限の任意の点 (x_1, x_2) において，$k>0$ として，$G(x_1, x_2)=[1-(1+x_1)^{-k}-(1+x_2)^{-k}+(1+x_1+x_2)^{-k}]$ で，その他の場合は 0 である $G(x_1, x_2)$ は 2 次元連続型確率変数 (x_1, x_2) の c.d.f に対するすべての条件を満足することを示し，その p.d.f. を求めよ．

2.7　(x, y) は連続型確率変数の対で，$x>0$，$y>0$ に対して p.d.f. が $f(x, y)$ となり，その他の場合は 0 であれば，次の各式を示せ．

（a）　$u=y/x$ の p.d.f. は

$$\int_0^\infty xf(x, ux)\,dx.$$

（b）　$v=x+y$ の p.d.f. は

$$\int_0^v f(x, v-x)\,dx.$$

（c）　$w=xy$ の p.d.f. は

$$\int_0^\infty \frac{1}{x}f\left(x, \frac{w}{x}\right)dx.$$

2.8　確率変数 (x_1, x_2) が x_1x_2 平面の各点 (x_1, x_2) で p.d.f. $(2\pi)^{-1}e^{-\frac{1}{2}(x_1^2+x_2^2)}$ を持つとする．y_1，y_2 が次に示すような x_1 と x_2 に関する確率変数ならば，y_1 と y_2 は独立であることを示せ．

$$x_1=y_1\cos y_2, \qquad x_2=y_1\sin y_2.$$

ただし y_1 の p.d.f. は $y_1>0$ に対して $y_1e^{-\frac{1}{2}y_1^2}$，$y_1\leqq 0$ に対して 0，そして y_2 の p.d.f. は $0<y_2<2\pi$ に対して $1/(2\pi)$ で，その他の場合は 0，となる．

2.9　3 次元確率変数 (x_1, x_2, x_3) が頂点 $(0, 0, 0)$，$(0, 0, 1)$，$(0, 1, 0)$，$(1, 0, 0)$ を持つ 4 面体の中で p.d.f. 6 を持ち，その他では 0 である．この場合，(x_1, x_2)，x_1，$x_1|x_2$，x_1+x_2，$x_1+x_2+x_3$ のおのおのの c.d.f. と p.d.f. を求めよ．

2.10　(x_1, x_2, x_3) が p.d.f. $e^{-(x_1+x_2+x_3)}$ を持つ非負の確率変数であれば，$u=x_1+x_2+x_3$ での (u, x_2, x_3) の p.d.f. を求めよ．(u, x_2, x_3) の p.d.f. から次の p.d.f. を求めよ．

（a）u　　（b）$u|x_2$　　（c）$u|x_2, x_3$　　（d）$(x_2, x_3|u)$.

2.11　$(x_1, \cdots, x_k)$ が c.d.f. $F(x_1, \cdots, x_k)$ を持つ独立な確率変数であれば，(2.6.4) 式で

$$\Delta_{I_k}^k F(x_1, \cdots, x_k)=\Delta F_1(x_1)\cdots\Delta F_k(x_k).$$

ただし

$$\Delta F_i(x_i)=F_i(x_i'')-F_i(x_i')$$

となることを示せ．また $F_i(x_i)$ が x_i の周辺 c.d.f. となることも示せ．

2.12　$(x_1, \cdots, x_k)$ が $x_1, \cdots, x_k$ の対称な p.d.f. を持つ k 次元確率変数であれば，$P(x_1<x_2<\cdots<x_k)=1/k!$ となることを示せ．

2.13　k 次元確率変数 $(x_1, \cdots, x_k)$ が $x_1>0, \cdots, x_k>0$ に対して次のような形の p.d.f. を持っていることがわかっている．

$$C(A+x_1+\cdots+x_k)^{-k-B}.$$

ただし A，B は正である．そしてその他の場合は 0 である．この場合の C を求めよ．また $r<k$ に対して $(x_1, \cdots, x_r)$ の周辺分布の p.d.f. を求めよ．

2.14　$(x_1, \cdots, x_k)$ が同一の連続型 c.d.f. $F(x_1), \cdots, F(x_k)$ を持つ独立な確率変数で，かつ

$$u=\min(x_1, \cdots, x_k), \qquad v=\max(x_1, \cdots, x_k)$$

であれば，(u, v) の c.d.f. は次のようになることを示せ．

$$[F(v)]^k-g(u, v)[F(v)-F(u)]^k.$$

ただし

$$g(u, v)=1, \quad u<v, \qquad g(u, v)=0, \quad u\geqq v.$$

また u と v の周辺 c.d.f. を求めよ．また $F(x)$ が微係数 $f(x)$ を持つとき，(u, v)，u，v の p.d.f. を求めよ．

2.15　壺の中に $1, 2, \cdots, N$ の数字をかいた N 枚の札がはいっている．1 枚の札を取り出し，その数を確率変数 x_1 で表わす．そして札はもとに戻す．次に第 2 番目の札を取り出し，その数を確率変数 x_2 で表わし，札をもとに戻す．これを k 回繰り返えす．各段階においてすべての N 個の数に等確率が与えてあり，$x_1, \cdots, x_k$ が互いに独立であるとき，k 次元確率変数 $(x_1, \cdots, x_k)$ の c.d.f. を求めよ．また $y+1$ が最小の数，すなわち引き出したどの数もこれより大きいとき，y の c.d.f. と p.f. を求めよ．

2.16（続き）　k $(<N)$ 個の札をもとに戻さずに連続して取り出したとする．ある段階で壺の中に残っている任意の札を取り出す確率は，その段階で残っている任意の他の札を

取り出すのと等しくなるようになっている.

$(x_1, \cdots, x_k)$ はいくつの質点を持つか？ また $(x_1, \cdots, x_k)$ の p.f. を定めよ. x の r 個 $(r < k)$ の任意の集合の周辺分布の p.f. は, x の r 個の任意の他の集合と同じになることを示せ.

2.17 $(x_1, y_1), \cdots, (x_k, y_k)$, $k \geqslant 2$ はそれぞれ同一の c.d.f. $F(x_1, y_1), \cdots, F(x_k, y_k)$ (同一の p.d.f. $f(x_1, y_1), \cdots, f(x_k, y_k)$ を持っている) を持つ独立な 2 次元確率変数であり, かつ

$$u = \max(x_1, \cdots, x_k), \qquad v = \max(y_1, \cdots, y_k)$$

ならば, (u, v) の p.d.f. が次のようになることを示せ.

$$k(k-1)F^{k-2}(u, v)\frac{\partial F}{\partial u}\cdot\frac{\partial F}{\partial v} + kF^{k-1}(u, v)f(u, v).$$

2.18 (続き) $w = \max(x_1 + y_1, \cdots, x_k + y_k)$ ならば, w の p.d.f. は次のようになることを示せ.

$$k\left[\int_{-\infty}^{\infty}\int_{-\infty}^{w-x} f(x, y)\,dy\,dx\right]^{k-1}\int_{-\infty}^{\infty} f(x, w-x)\,dx.$$

2.19 確率変数 $x_1, \cdots, x_k$ が 1 つを残して独立であれば, 互いに独立であることを示せ.

2.20 $F(x_1, \cdots, x_k)$ が周辺 c.d.f. $F_1(x_1)\cdots F_k(x_k)$ を持つ k 次元 c.d.f. であれば

$$F(x_1, \cdots, x_k) \leqslant [F_1(x_1)\cdots F_k(x_k)]^{1/k}$$

となることを示せ.

第3章　確率変数の平均値とモーメント

3.1　概　説

1.8 節で, 確率空間 $(R, \mathscr{B}, P)$ の集合 $E \in \mathscr{B}$ 上での確率変数 $x(e)$ の P に関するルベーグ=スティルチェス積分を定義した.

ここで E' を R_1 におけるボレル集合, E を $x(e) \in E'$ なる R の集合として, $F(x)$ が $x(e)$ の c.d.f. ならば, E 上の $x(e)$ のルベーグ=スティルチェス積分 (1.8.5) を次の等式で表わす.

(3.1.1)
$$\int_E x(e)\,dP(e) = \int_{E'} x\,dF(x).$$

同じく E'_k を R_k のボレル集合, E を $(x_1(e), \cdots, x_k(e)) \in E'_k$ となる R の集合とする. $F(x_1, \cdots, x_k)$ が $(x_1(e), \cdots, x_k(e))$ の c.d.f. で, $g(x_1, \cdots, x_k)$ が $\mathscr{B}_k$ に関して可測ならば [したがって $g(x_1(e), \cdots, x_k(e))$ は $\mathscr{B}$ に関して可測となる], 次のように書ける.

(3.1.2)
$$\int_E g(x_1(e), \cdots, x_k(e))\,dP(e) = \int_{E'_k} g(x_1, \cdots, x_k)\,dF(x_1, \cdots, x_k).$$

E'_1 が R_1 のボレル集合で, (3.1.2) の E'_k が $g(x_1, \cdots, x_k) \in E'_1$ となる R_k の点集合であるとし, $H(y)$ が $g(x_1, \cdots, x_k)$ の c.d.f. であれば, 次のように書くことができる.

(3.1.3)
$$\int_{E'_k} g(x_1, \cdots, x_k)\,dF(x_1, \cdots, x_k) = \int_{E'_1} y\,dH(y).$$

3.2　平　均　値

(3.1.1) で R_1 を集合 E' にすれば，その結果から生じる積分は確率変数 x の平均値を定義している．すなわち次のように書く．

(3.2.1) $$\mathscr{E}(x)=\int_{-\infty}^{\infty}x\,dF(x).$$

x が 2.3 (a) 節で定義した p.f. $p(x)$ を持つ離散型確率変数であれば，$\mathscr{E}(x)$ は次のような和になる．

(3.2.2) $$\mathscr{E}(x)=\sum_{\alpha}x^{(\alpha)}p(x^{(\alpha)}).$$

ただし $x^{(\alpha)}$ は確率変数 x の質点である．

x が p.d.f. $f(x)$ を持つ連続型確率変数ならば

(3.2.3) $$\mathscr{E}(x)=\int_{-\infty}^{\infty}x\,dF(x)=\int_{-\infty}^{\infty}xf(x)\,dx.$$

一般に，(3.1.2) で E'_k を全空間 R_k にすれば

(3.2.4) $$\mathscr{E}(g(x_1,\cdots,x_k))=\int_{R_k}g(x_1,\cdots,x_k)\,dF(x_1,\cdots,x_k)$$

が得られる．さらに (3.1.3) から次のことがわかる．$E'_k=R_k$ ならば $E'_1=R_1$ で

(3.2.5) $$\mathscr{E}(g(x_1,\cdots,x_k))=\int_{-\infty}^{\infty}y\,dH(y)$$

となる．

$g(x_1,\cdots,x_k)$ を y で表わせば，(3.2.4) の積分は単に $\mathscr{E}(y)$ となり，次のようになる．

(3.2.6) $$\mathscr{E}(g(x_1,\cdots,x_k))=\mathscr{E}(y).$$

$(x_1,\cdots,x_k)$ が離散型確率変数であれば，(3.2.4) は和となり，$(x_1,\cdots,x_k)$ が連続型であれば，(3.2.4) は R_k 上の k 次元積分となる．

$g(x_1,\cdots,x_k)$ が確率変数で，モーメントとして y を用いれば，平均値は，それが存在するとき，次のような性質を持つ．

3.2.1　$\mathscr{E}(cy)=c\mathscr{E}(y)$　　ただし，c は定数．

3.2.2　$\mathscr{E}(ay_1+by_2)=a\mathscr{E}(y_1)+b\mathscr{E}(y_2)$.

3.2.3　$m\leqslant y\leqslant M$　ならば　$m\leqslant\mathscr{E}(y)\leqslant M$.

3.2.4　$\mathscr{E}(y_1)\leqslant\mathscr{E}(y_2)$　　ただし，$y_1\leqslant y_2$.

3.2.5　$|\mathscr{E}(y)|\leqslant\mathscr{E}(|y|)$.

(x_1,x_2) を x_1 と x_2 が独立で，c.d.f. $F(x_1,x_2)$ を持つ確率変数として，積 x_1x_2 の平均値を考えてみる．次のことが証明できる．

(3.2.7) $$\mathscr{E}(x_1x_2)=\int_{-\infty}^{\infty}x_1\,dF_1(x_1)\cdot\int_{-\infty}^{\infty}x_2\,dF_2(x_2)=\mathscr{E}(x_1)\cdot\mathscr{E}(x_2).$$

したがって次の結果が得られる．

3.2.6　(x_1,x_2) が確率変数で，その成分 x_1 と x_2 が独立であれば
$$\mathscr{E}(x_1x_2)=\mathscr{E}(x_1)\cdot\mathscr{E}(x_2).$$

もちろん，k 個の互いに独立な確率変数の場合にも同じ結果が得られる．

3.3　1変数のモーメント

数理統計学には，確率変数の c.d.f. を完全に定めることが不可能か，あるいは不可能に近い場合が多い．このような場合，しばしば確率変数の分布を確率変数のモーメントとモーメントのある種の関数によって，有効に記述することが可能である．

x を c.d.f. $F(x)$ を持つ確率変数とする．x の平均値は (3.2.1) で定義したように通常 $\mathscr{E}(x)$ または $\mu(x)$ で表わされる．すなわち

(3.3.1) $$\mathscr{E}(x)=\mu(x)=\int_{-\infty}^{\infty}x\,dF(x).$$

x の分散[*)]は確率変数 $(x-\mu(x))^2$ の平均値として定義される．すなわち

(3.3.2) $$\sigma^2(x)=\mathscr{E}(x-\mu(x))^2.$$

$\sigma^2(x)$ の正の平方根 $\sigma(x)$ は x の標準偏差と呼ばれる．比 $[\sigma(x)]/\mu(x)$ を x の変動係数という．

力学の概念と用語を用いると，確率変数 x の平均は x の確率分布の R_1 における重心と解釈できる．x の分散は，その重心のまわりの同じ確率分布の慣性率と解釈できるし，

[*)] $\mathscr{E}(x)$ のかわりに ave (x) を，$\sigma(x)^2$ のかわりに var (x) を用いて表わすこともある．

また，確率質点が重心のまわりに広がる（または集まる）量を表わすものでもある．

確率変数がわかっている場合には $\mu(x)$ を μ と書き，$\sigma(x)$ は σ と書く．

(3.3.1) と **3.2.1** および **3.2.2** で示した平均値の性質を用いれば，次の式が得られる．

(3.3.3) $$\sigma^2 = \mathscr{E}(x^2 - 2x\mu + \mu^2) = \mathscr{E}(x^2) - \mu^2.$$

a が任意の定数のとき，平均値 $\mathscr{E}(x-a)^2$ をとると次のようになる．

(3.3.4) $$\mathscr{E}(x-a)^2 = \mathscr{E}[(x-\mu)+(\mu-a)]^2 = \sigma^2 + (\mu-a)^2.$$

これから次のことがいえる．

3.3.1 $\mathscr{E}(x-a)^2$ を最小にする定数 a の値は μ であり，$\mathscr{E}(x-a)^2$ の最少値は σ^2 である．

確率分布の“尾に当る部分”の確率は，次に述べる **Chebyshev** (1867) の不等式で与えられる．

3.3.2 x は確率変数で μ をその平均，$\sigma^2 > 0$ を分散とすれば

(3.3.5) $$P(|x-\mu| \geqslant \lambda\sigma) \leqslant \frac{1}{\lambda^2}.$$

ただし λ は任意の正の定数とする．

これを証明するため，まず x 軸 R_1 を 3 つの互いに素な区間に分割する．

$$I = (-\infty, \mu-\lambda\sigma], \qquad I' = (\mu-\lambda\sigma, \mu+\lambda\sigma), \qquad I'' = [\mu+\lambda\sigma, +\infty).$$

すると x の分散を次のように書くことができる．

(3.3.6) $$\sigma^2 = \int_I (x-\mu)^2 dF(x) + \int_{I'} (x-\mu)^2 dF(x) + \int_{I''} (x-\mu)^2 dF(x).$$

(3.3.6) の右辺の中央の項をのぞけば，次の式が得られる．

(3.3.7) $$\sigma^2 \geqslant \int_I (x-\mu)^2 dF(x) + \int_{I''} (x-\mu)^2 dF(x).$$

右辺の最初の積分で $(x-\mu)^2$ の x を $\mu-\lambda\sigma$ で置き換え，2 番目の項の $(x-\mu)^2$ における x を $\mu+\lambda\sigma$ で置き換えても不等式はそのまま成立するから

(3.3.8) $$\sigma^2 \geqslant \sigma^2\lambda^2 P(|x-\mu| \geqslant \lambda\sigma).$$

これは不等式 (3.3.5) と同じである．

r を正の整数とすると r 次のモーメント $\mu_r'(x)$ は，存在すれば確率変数 x^r の平均値で定義される．すなわち

(3.3.9) $$\mu_r'(x) = \mathscr{E}(x^r).$$

便宜上，$r=0$ の場合には $\mu_0'(x) = 1$ とする．$\mu_r'(x)$ が存在するか，有限か，という問題は，単に確率変数 x^r が R_1 上で積分可能か否かというのと同じことである．r 次の中心モーメント $\mu_r(x)$ は $(x-\mu)^r$ の平均値として定義される．すなわち

(3.3.10) $$\mu_r(x) = \mathscr{E}[(x-\mu)^r].$$

確率変数が明らかなときには，$\mu_r'(x)$ および $\mu_r(x)$ を μ_r', μ_r で表わす．

μ_r, $r = 1, 2, \cdots$ が μ_r' の多項式で表現でき，またその逆も可能である．$\mu_1' = \mu$ に注意すれば次の式が得られる．

(3.3.11) $$\begin{aligned} \mu_1 &= 0 \\ \mu_2 &= \mu_2' - \mu^2 \\ \mu_3 &= \mu_3' - 3\mu_2'\mu + 2\mu^3 \\ &\vdots \end{aligned}$$

逆に

(3.3.12) $$\begin{aligned} \mu_1' &= \mu \\ \mu_2' &= \mu_2 + \mu^2 \\ \mu_3' &= \mu_3 + 3\mu_2\mu + \mu^3 \\ &\vdots \end{aligned}$$

ここで $\mu_2 = \sigma^2$ である．

r 次の絶対モーメント $\nu_r'(x)$ は次のように定義される．

(3.3.13) $$\nu_r'(x) = \mathscr{E}(|x|^r).$$

通常 $\nu_r'(x)$ を ν_r' と書く．r 次の絶対中心モーメントを次のように定義する．

(3.3.13 a) $$\nu_r(x) = \mathscr{E}(|x-\mu|^r).$$

一般に，任意の確率変数 $g(x)$ の分散と高次モーメントは x の c.d.f. で表わすことができる．たとえば，(3.3.1) と (3.3.2) を用いれば，$g(x)$ の平均と分散は次のように定義される．

(3.3.14) $$\mu(g(x)) = \mathscr{E}[g(x)] \quad \text{および} \quad \sigma^2(g(x)) = \mathscr{E}\{(g(x))^2 - [\mu(g(x))]^2\}.$$

離散型確率変数を含む問題では，まず階乗モーメントの値を求めることによって，モーメント μ_r' を定めると便利である．

(3.3.15) $$x^{[r]} = x(x-1)\cdots(x-r+1)$$

とすれば r 次の階乗モーメントは次のように定義される．

(3.3.16) $$\mu_{[r]}'(x) = \mathscr{E}(x^{[r]}).$$

そして変数が明らかなときは $\mu'_{[r]}(x)$ を $\mu'_{[r]}$ と書く．これをモーメントを用いて表わせば次のようになる．

$$
(3.3.17)\qquad
\begin{aligned}
\mu'_{[1]} &= \mu'_1\\
\mu'_{[2]} &= \mu'_2 - \mu'_1\\
\mu'_{[3]} &= \mu'_3 - 3\mu'_2 + 2\mu'_1\\
&\vdots
\end{aligned}
$$

また逆に

$$
(3.3.18)\qquad
\begin{aligned}
\mu'_1 &= \mu'_{[1]}\\
\mu'_2 &= \mu'_{[2]} + \mu'_{[1]}\\
\mu'_3 &= \mu'_{[3]} + 3\mu'_{[2]} + \mu'_{[1]}\\
&\vdots
\end{aligned}
$$

これまでに“確率変数 x のモーメント”として種々のモーメントを定義してきた．それらは“x の確率分布のモーメント”を指すこともあるし，$F(x)$ が x の c.d.f. であれば，“$F(x)$ のモーメント”を指すこともある．

3.4　2 変数のモーメント

(x_1, x_2) を c.d.f. $F(x_1, x_2)$ を持つ 2 次元確率変数とする．x_1 と x_2 の平均値は

$$(3.4.1)\qquad \mu(x_1) = \mathscr{E}(x_1), \qquad \mu(x_2) = \mathscr{E}(x_2)$$

となり，普通 μ_1, μ_2 で表わす．分散は

$$(3.4.2)\qquad \sigma^2(x_1) = \mathscr{E}(x_1 - \mu_1)^2, \qquad \sigma^2(x_2) = \mathscr{E}(x_2 - \mu_2)^2$$

となり，σ_1^2, σ_2^2 で表わす．

$\mu(x_1)$ を μ_1 で，また $\mu(x_2)$ を μ_2 で表わすと (3.3.10) と (3.4.1) のちがいが不明確になる．しかしさほど問題ではない．それは (3.3.10) の意味が $\mu_1 = 0$, $\mu_2 = \mu'_2 - \mu^2$ であり，これらの記号が現われる場合はいつもはっきりと表示されているからである．

x_1 と x_2 の**共分散**を $\mathrm{cov}\,(x_1, x_2)$ と書き，次のように定義する．

$$(3.4.3)\qquad \mathrm{cov}\,(x_1, x_2) = \mathscr{E}[(x_1 - \mu_1)(x_2 - \mu_2)].$$

簡単にすると次のようになる．

$$(3.4.4)\qquad \mathrm{cov}\,(x_1, x_2) = \mathscr{E}(x_1 x_2) - \mathscr{E}(x_1)\cdot\mathscr{E}(x_2).$$

$\mathrm{cov}\,(x_1, x_2) = \mathrm{cov}\,(x_2, x_1)$ であることに注意しよう．

x_1 と x_2 が独立ならば **3.2.6** から，$\mathscr{E}(x_1, x_2) = \mathscr{E}(x_1)\cdot\mathscr{E}(x_2)$ となり，次の結果が得られる．

3.4.1　x_1 と x_2 が独立ならば

$$(3.4.5)\qquad \mathrm{cov}\,(x_1, x_2) = 0.$$

注意　$\mathrm{cov}\,(x_1, x_2) = 0$ という条件は，x_1 と x_2 が独立であることを意味するとは限らない．たとえば，(x_1, x_2) が確率 1 で $x_2 = \cos x_1$ となる確率変数とし，x_1 の p.d.f. が次の式で与えられるとする．

$$f_1(x_1) = \begin{cases} \dfrac{1}{2\pi}, & -\pi < x_1 < +\pi \\ 0, & \text{その他}. \end{cases}$$

$\mathrm{cov}\,(x_1, x_2) = 0$ となるが，x_2 は確率 1 で x_1 から完全に（関数として）定められる．

x_1 と x_2 の間の**相関係数** $\rho(x_1, x_2)$ は

$$(3.4.6)\qquad \rho(x_1, x_2) = \frac{\mathrm{cov}\,(x_1, x_2)}{\sigma_1\sigma_2}$$

で定義され，確率変数を明示する必要がなければ，ρ_{12} または ρ と書く．$\rho(x_1, x_2) = 0$ であれば，x_1 と x_2 は**相関がない**という．

t を実数の定数とし，次の確率変数の平均値をとる．

$$\left[t\left(\frac{x_1 - \mu_1}{\sigma_1}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)\right]^2.$$

すると次のようになる．

$$(3.4.7)\qquad \mathscr{E}\left[t\left(\frac{x_1 - \mu_1}{\sigma_1}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)\right]^2 = t^2 + 2t\rho + 1 \geqslant 0.$$

すべての実数 t に対して，$t^2 + 2t\rho + 1 \geqslant 0$ となるための条件は $\rho^2 - 1 \leqslant 0$ である．したがって

3.4.2　相関係数 ρ は条件 $-1 \leqslant \rho \leqslant 1$ を満足する．

x_1 と x_2 が真に 1 次従属 [2.8 (b) 節を見よ] であれば $P(x_2 = \beta_0 + \beta_1 x_1) = 1$，ただし $\beta_1 \neq 0$．これから $\sigma^2(x_2 - \beta_0 - \beta_1 x_1) = 0$ となり，$\rho = +1$ または -1 となるのは β_1 が正の数かまたは負の数により定まる．逆に $\rho = \pm 1$ ならば，確率変数

(3.4.8)
$$\mp\left(\frac{x_1-\mu_1}{\sigma_1}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)$$

は平均値が 0 で退化する．これは $\rho=\pm 1$ ならば

(3.4.9)
$$P\left(\mp\left(\frac{x_1-\mu_1}{\sigma_1}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)=0\right)=1$$

となることを意味する．

したがって

3.4.3　有限な分散を持つ退化しない確率変数 x_1 と x_2 が（真に）1 次従属であるための必要十分条件は $\rho^2=1$ である．

モーメント $\mu'_{r_1r_2}$ と中心モーメント $\mu_{r_1r_2}$ は次のように定義される．

(3.4.10)
$$\mu'_{r_1r_2}=\mathscr{E}(x_1^{r_1}x_2^{r_2})$$

(3.4.11)
$$\mu_{r_1r_2}=\mathscr{E}[(x_1-\mu_1)^{r_1}(x_2-\mu_2)^{r_2}].$$

この表記法では，x_1 と x_2 の平均は μ'_{10} と μ'_{01} であり，x_1 と x_2 の分散はそれぞれ μ_{20} と μ_{02} である．x_1 と x_2 の共分散は μ_{11} であり，相関係数は $\rho_{12}=\mu_{11}/\sqrt{\mu_{20}\mu_{02}}$ である．

絶対モーメント $\nu'_{r_1r_2}$ と階乗モーメント $\mu'_{[r_1][r_2]}$ は次のように定義される．

(3.4.12)
$$\nu'_{r_1r_2}=\mathscr{E}(|x_1|^{r_1}\cdot|x_2|^{r_2})$$

(3.4.13)
$$\nu_{r_1r_2}=\mathscr{E}(|x_1-\mu_1|^{r_1}\cdot|x_2-\mu_2|^{r_2})$$

(3.4.14)
$$\mu'_{[r_1][r_2]}=\mathscr{E}(x_1^{[r_1]}x_2^{[r_2]}).$$

3.5　k 変数のモーメント

前述の定義の k 次元確率変数の場合は，そのまま拡張すれば良い．すなわち $(x_1,\cdots,x_k)$ が c.d.f. $F(x_1,\cdots,x_k)$ を持つ k 次元確率変数であれば，x_i の平均値は (3.2.6) の $g(x_1,\cdots,x_k)=x_i$ で与えられ，$H(y)$ が x_i の周辺 c.d.f. のとき (3.2.5) となる．簡単にいえば

(3.5.1)
$$\mu_i=\mathscr{E}(x_i)=\int_{R_k}x_i\,dF(x_1,\cdots,x_k)=\int_{-\infty}^{\infty}x_i\,dF_i(x_i).$$

x_i の分散 σ_i^2 は定義 (3.3.2) に従って，x_i の周辺 c.d.f. から同様に定義される．また，x_i と x_j の共分散は定義 (3.4.3) に従って，周辺 c.d.f. $F_{ij}(x_i,x_j)$ から定義さ

れる．

確率変数 $(x_1,\cdots,x_k)$ の k 個の成分のすべての分散と共分散の集合は次の $k\times k$ の対称行列をつくり，**共分散行列**と呼ばれる．

(3.5.2)
$$\|\sigma_{ij}\|,\qquad i,j=1,\cdots,k.$$

σ_{ii} は x_i の分散で，σ_{ij}, $i\neq j$ は x_i と x_j の共分散である．確率変数を明示するときには，$\|\sigma(x_i,x_j)\|$ と書く．

便宜上，確率変数 $(x_1,\cdots,x_k)$ が平均 $(\mu_1,\cdots,\mu_k)$ と共分散行列 $\|\sigma_{ij}\|$ を持つという．

(3.5.2) の共分散行列だけでなく，**共分散行列の逆行列**もまた有用である．すなわち $\|\sigma_{ij}\|^{-1}$ で次のように書かれる．

(3.5.3)
$$\|\sigma^{ij}\|.$$

ただし σ^{ij} は正式に表わせば，$\sigma^{ij}=(\|\sigma_{ij}\|$ の σ_{ij} の余因子$)/|\sigma_{ij}|$，となり，$|\sigma_{ij}|$ は行列 $\|\sigma_{ij}\|$ の行列式である．$\|\sigma^{ij}\|$ は $|\sigma_{ij}|\neq 0$ のときにのみ存在する．また $\|\sigma_{ij}\|$ が逆行列 $\|\sigma^{ij}\|$ を持つならば $|\sigma^{ij}|=\dfrac{1}{|\sigma_{ij}|}$ である．**相関行列** $\|\rho_{ij}\|$ に対しても同様なことがいえる．

3.4.3 と同様に，退化しない確率変数 x_1 と x_2 が真に 1 次従属であるための必要十分条件は $\rho^2=1$ である．それと同値な必要十分条件を次のように述べることができる．

$$\begin{vmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{vmatrix}=0.$$

k 次元確率変数の場合にも，1 次従属となる基準があると便利である．2.8 (d) 節で述べたように，すべて 0 でない定数の組 $c_1,\cdots,c_k$ が存在し，$c_1x_1+\cdots+c_kx_k$ が退化する確率変数のとき，$(x_1,\cdots,x_k)$ の成分は真に 1 次従属である．真に 1 次従属であるための条件は定理 3.4.3 の拡張で与えられる．

3.5.1　確率変数 $(x_1,\cdots,x_k)$ の成分が真に 1 次従属であるための必要十分条件は，行列 $\|\sigma_{ij}\|$ の階数が $k-1$ となることである．

まず十分性を考える．

$\|\sigma_{ij}\|$ の階数は $k-1$ であるから $|\sigma_{ij}|=0$ であり，また $|\sigma_{ij}|$ のすべての主小行列式は正である．したがって，すべて 0 でない実数の定数の組 c_i, $i=1,\cdots,k$ が存在して，$\sum_{i=1}^{k}\sigma_{ij}c_i=0$, $j=1,\cdots,k$ となる．この等式に c_j をかけて j に関して和をとれば，次の式が得られる．

(3.5.4)
$$\sum_{i,j=1}^{k} \sigma_{ij}c_i c_j = 0.$$

これは次のようにも書ける．

(3.5.5)
$$\mathscr{E}\left[\sum_{i=1}^{k}(x_i-\mu_i)c_i\right]^2 = 0.$$

これは確率変数 $\sum_{i=1}^{k} c_i x_i$ の分散である．分散が 0 であるから $\sum_{i=1}^{k} c_i x_i$ は退化する．そしてどの c_i も 0 ではないから，定義より $x_1, \cdots, x_k$ は真に 1 次従属である．

次に必要性を示そう．$x_1, \cdots, x_k$ が真に 1 次従属であれば，定義によって退化する確率変数 $\sum_{i=1}^{k} x_i c_i$ が存在する．ただし c_i はすべて 0 でない実数の定数である．したがって

(3.5.6)
$$\sigma^2\left(\sum_{i=1}^{k} c_i x_i\right) = \sum_{i,j=1}^{k} \sigma_{ij}c_i c_j = 0.$$

さて $\phi(c_1, \cdots, c_k) = \sum_{i,j=1}^{k} \sigma_{ij}c_i c_j$ を c_i の関数として考える．$\phi(c_1, \cdots, c_k)$ は負にはなりえないから，仮定より，どの c_i も 0 でないような点（c の空間内の）で 0 なる最小値をとる．しかし，$\phi(c_1, \cdots, c_k)$ が最小値をとる c_i の値は，次の等式を満たす．

(3.5.7)
$$\frac{\partial\phi}{\partial c_i} = 0, \qquad i = 1, \cdots, k.$$

この等式は次のように変形できる．

(3.5.8)
$$\sum_{j=1}^{k} \sigma_{ij}c_j = 0, \qquad i = 1, \cdots, k.$$

各 e_i は 0 でないから $|\sigma_{ij}| = 0$ でなければならない．$|\sigma_{ij}|$ のどれかの主小行列式が 0 となれば，c がすべて 0 でないという仮定に矛盾する．したがって $\|\sigma_{ij}\|$ の階数は $k-1$ であり，これで **3.5.1** の証明は終わった．

実数 $c_1, \cdots, c_k$ に対して 2 次形式 $\sum_{i,j=1}^{k} \sigma_{ij}c_i c_j$ が消滅する（すなわち $c_1 = \cdots = c_k = 0$ のときにのみ最小値 0 をとる）ならば，この 2 次形式は**正定値**[*)] であるといい，その行列 $\|\sigma_{ij}\|$ を**正定値行列**という．共分散行列の正定値に関する定理を次に述べる．

3.5.2　$(x_1, \cdots, x_k)$ が退化しない成分を持つ k 次元確率変数で，共分散行列 $\|\sigma_{ij}\|$ を持つとき，$\|\sigma_{ij}\|$ が正定値であるための必要十分条件は，成分 $x_1, \cdots, x_k$ の間に 1 次従属なものが存在しないことである．

*)　正の定符号ともいう．（訳注）

この証明は読者にゆだねる．

2 次形式に関する次の性質は後節でも用いる．

3.5.3　$\sum_{i,j=1}^{k} \sigma_{ij}c_i c_j$ が正定値ならば，2 次形式 $\sum_{i,j=1}^{k} \sigma^{ij}c_i c_j$ もまた正定値である．

この証明は簡単なので省略する．

k 次元確率変数 $(x_1, \cdots, x_k)$ に関する高次モーメント $\mu'_{r_1\cdots r_k}$，$\mu_{r_1\cdots r_k}$，$\nu'_{r_1\cdots r_k}$，$\nu_{r_1\cdots r_k}$ および $\mu'_{[r_1]\cdots[r_k]}$ は (3.4.10)，(3.4.11)，(3.4.12)，(3.4.13)，(3.4.14) の拡張によって定義される．

3.6　線形関数の平均，分散，共分散

いくつかの確率変数の関数で頻度が高く現われる形は線形関数である．この関数の平均，分散について次に示す．

3.6.1　$(x_1, \cdots, x_k)$ が平均 $(\mu_1, \cdots, \mu_k)$ と共分散行列 $\|\sigma_{ij}\|$ を持つ k 次元確率変数であれば，$c_1, \cdots, c_k$ を定数とすると，

(3.6.1)
$$L = \sum_{i=1}^{k} c_i x_i$$

なる線形関数の平均および分散は

(3.6.2)
$$\mathscr{E}(L) = \sum_{i=1}^{k} c_i \mu_i$$

(3.6.3)
$$\sigma^2(L) = \sum_{i,j=1}^{k} \sigma_{ij}c_i c_j.$$

証明は読者にゆだねる．

x_i が独立な場合には，**3.4.1** から $\sigma_{ij} = 0$，$i \neq j$ となり，**3.6.1** の系が得られる．

3.6.1 a　$x_1, \cdots, x_k$ が相関のない確率変数で，分散 $\sigma_1^2, \cdots, \sigma_k^2$ を持つならば，$L\left(= \sum_{i=1}^{k} c_i x_i\right)$ の分散は

(3.6.4)　$$\sigma^2(L)=\sum_{i=1}^{k}c_i^2\sigma_i^2.$$

一般には

3.6.2　$L_p=\sum_{i=1}^{k}c_{ip}x_i,\quad p=1,\cdots,s$ を **3.6.1** で述べた確率変数の線形関数とすれば，L_p は平均

(3.6.5)　$$\mathscr{E}(L_p)=\sum_{i=1}^{k}c_{ip}\mu_i,\qquad p=1,\cdots,s$$

と共分散行列は

(3.6.6)　$$\|\sigma(L_p,L_q)\|=\left\|\sum_{i,j=1}^{k}\sigma_{ij}c_{ip}c_{jq}\right\|$$

となる．

$x_1,\cdots,x_k$ が分散 $\sigma_1^2,\cdots,\sigma_k^2$ を持つ相関のない確率変数であれば，共分散行列 $\|\sigma(L_p,L_q)\|$ は $\left\|\sum_{i=1}^{k}\sigma_i^2c_{ip}c_{iq}\right\|$ となる．

3.7　条件つき確率変数の平均値

(a)　2 変数の場合

$x_2|x_1$ を，2.9 (b) 節で定義したように，c.d.f. が $F(x_2|x_1)$ である条件つき確率変数とする．$x_2|x_1$ の平均値を，存在すれば，次のように定義する．

(3.7.1)　$$\mu(x_2|x_1)=\mathscr{E}(x_2|x_1)=\int_{-\infty}^{\infty}x_2\,dF(x_2|x_1).$$

$\mu(x_2|x_1)$ は，x_1 の関数と考えられるが，x_1 上への x_2 の**回帰関数**と呼ばれる．図形的には x_1 の関数として，条件つき確率変数 $x_2|x_1$ の重心の軌跡を表わしている．特に，

(3.7.2)　$$\mu(x_2|x_1)=\beta_0+\beta_1x_1$$

ならば，x_2 の x_1 上の **線形回帰関数**が得られ，β_0 と β_1 は回帰係数と呼ばれる．

一般的には，条件つき確率変数 $g(x_2)|x_1$ の平均値は次のように表わされる．

(3.7.3)　$$\mathscr{E}(g(x_2)|x_1)=\int_{-\infty}^{\infty}g(x_2)\,dF(x_2|x_1).$$

特に $x_2|x_1$ の分散は次で与えられる．

(3.7.4)　$$\sigma^2(x_2|x_1)=\mathscr{E}[(x_2-\mu(x_2|x_1))^2|x_1].$$

$\mu(x_2|x_1)$，$\sigma^2(x_2|x_1)$ は与えられた x_1 のもとでの x_2 の**条件つき平均**および**条件つき分散**という．$\sigma^2(x_2|x_1)$ は x_2 の x_1 上への**残差分散**と呼ばれる．

3.3.1 の系として，$\sigma^2(x_2|x_1)$ と $\mu(x_2|x_1)$ との関係を次に示す．

3.7.1　$x_2|x_1$ が 2.9 (b) 節で定義した条件つき確率変数で，$u(x_1)$ が x_1 の実数値で 1 価の関数ならば，$\mathscr{E}[(x_2-u(x_1))^2|x_1]$ を最小にする $u(x_1)$ の値は $u(x_1)=\mu(x_2|x_1)$ で与えられる．さらに，$\mathscr{E}[(x_2-u(x_1))^2|x_1]$ の最小値は $\sigma^2(x_2|x_1)$ である．

$x_2|x_1$ が定義されているところで，すべての x_1 の値に対して $\sigma^2(x_2|x_1)=0$ ならば，明らかに $P[(x_2=\mu(x_2|x_1))|x_1]=1$ である．このことは，確率変数 x_1 の値が，ある1つの事象に対して既知であれば（または与えられれば），その事象に対する x_2 の値は確率 1 で $\mu(x_2|x_1)$ である，ということを意味している．

(3.7.3) で $g(x_2)=x_2^r$ とおけば，条件つき確率変数 $x_2|x_1$ の r 次のモーメントが得られる．$x_2|x_1$ の r 次の中心モーメントおよび階乗モーメントも同様に定義できる．

確率変数 (x_1,x_2) の関数 $g(x_1,x_2)$ の平均値を実際に求めるには，繰り返し積分を行なう．その場合，条件つき確率変数は重要な役割を果たす．次にそのことを示そう．証明は省く．

3.7.2　(x_1,x_2) を c.d.f. $F(x_1,x_2)$ を持つ確率変数とする．$x_2|x_1$ および $x_1|x_2$ を 2.9 (b) 節の意味での c.d.f. $F(x_2|x_1)$ および $F(x_1|x_2)$ を持つ条件つき確率変数とする．$g(x_1,x_2)$ が確率変数のとき

(3.7.5)　$$\begin{aligned}\int_{R_2}g(x_1,x_2)\,dF(x_1,x_2)&=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}g(x_1,x_2)\,dF(x_2|x_1)\right]dF_1(x_1)\\&=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}g(x_1,x_2)\,dF(x_1|x_2)\right]dF_2(x_2).\end{aligned}$$

もっと簡潔に

(3.7.5 a)　$$\begin{aligned}\mathscr{E}(g(x_1,x_2))&=\mathscr{E}_{(x_1)}[\mathscr{E}_{(x_2)}(g(x_1,x_2)|x_1)]\\&=\mathscr{E}_{(x_2)}[\mathscr{E}_{(x_1)}(g(x_1,x_2)|x_2)].\end{aligned}$$

2.9 節で述べた条件つき確率変数よりも簡単ではあるが，一般型として **3.7.2** は容易

に証明できよう．

ラドン=ニコディムの定理 **1.10.1** を用いれば，**3.7.1** および **3.7.2** をさらに一般的な条件つき確率変数の場合に拡張できる．

(b) 数個の変数の場合

$x_k|x_1,\cdots,x_{k-1}$ を 2.9 (c) 節の意味での c.d.f. $F(x_k|x_1,\cdots,x_{k-1})$ を持つ条件つき確率変数とする．条件つき確率変数 $g(x_k)|x_1,\cdots,x_{k-1}$ の平均値は次のように定義される．

(3.7.6)
$$\mathscr{E}(g(x_k)|x_1,\cdots,x_{k-1})=\int_{-\infty}^{\infty}g(x_k)\,dF(x_k|x_1,\cdots,x_{k-1}).$$

特に $x_k|x_1,\cdots,x_{k-1}$ の平均と分散は次のように定義される．

(3.7.7)
$$\mu(x_k|x_1,\cdots,x_{k-1})=\mathscr{E}(x_k|x_1,\cdots,x_{k-1})$$

および

(3.7.8)
$$\sigma^2(x_k|x_1,\cdots,x_{k-1})=\mathscr{E}[(x_k-\mu(x_k|x_1,\cdots,x_{k-1}))^2|x_1,\cdots,x_{k-1}].$$

読者は，**3.7.1** が確率変数 $x_k|x_1,\cdots,x_{k-1}$ の場合にただちに拡張できることに気づくだろう．

一般に，$x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1}$, $k=k_1+k_2$ が k_2 次元条件つき確率変数で，c.d.f. $F(x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1})$ を持つとき，確率変数 $g(x_{k_1+1},\cdots,x_k)|x_1,\cdots,x_{k_1}$ の平均値は

(3.7.9)
$$\mathscr{E}[g(x_{k_1+1},\cdots,x_k)|x_1,\cdots,x_{k_1}]$$
$$=\int_{R_{k_2}}g(x_{k_1+1},\cdots,x_k)\,dF(x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1}).$$

さて，条件つき確率変数 $x_k|x_1,\cdots,x_{k-2}$ と $x_{k-1}|x_1,\cdots,x_{k-2}$ の間の共分散や，結合モーメントなどの定義の仕方は明らかである．

$x_1,\cdots,x_{k-1}$ の関数としての，値 $\mu(x_k|x_1,\cdots,x_{k-1})$ を x_k の $x_1,\cdots,x_{k-1}$ 上への回帰関数と呼ぶ．次のような特別の場合には，x_k は $x_1,\cdots,x_{k-1}$ 上の線形回帰関数を持つ．

(3.7.10)
$$\mu(x_k|x_1,\cdots,x_{k-1})=\beta_0+\beta_1x_1+\cdots+\beta_{k-1}x_{k-1}.$$

最後に，**3.7.2** を拡張して，(3.7.5) を一般化したものが得られる．

(3.7.11)
$$\int_{R_k}g(x_1,\cdots,x_k)\,dF(x_1,\cdots,x_k)$$
$$=\int_{R_{k_1}}\left[\int_{R_{k_2}}g(x_1,\cdots,x_k)\,dF(x_{k_1+1},\cdots,x_k|x_1,\cdots,x_{k_1})\right]\cdot dF_{1\cdots k_1}(x_1,\cdots,x_{k_1}).$$

または，完全な積分の繰り返しとして書けば

(3.7.12)
$$\int_{R_k}g(x_1,\cdots,x_k)\,dF(x_1,\cdots,x_k)$$
$$=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,\cdots,x_k)\,dF(x_k|x_1,\cdots,x_{k-1})$$
$$\cdot dF_{1\cdots k-1}(x_{k-1}|x_1,\cdots,x_{k-2})\cdots dF_1(x_1).$$

もちろん，繰り返しの可能な順序としては $k!$ 個ある．

ラドン=ニコディムの定理を用いて，(3.7.6) から (3.7.12) までの式に，2.9 (c) 節での条件よりもさらに一般的な $F(x_1,\cdots,x_k)$ に関する条件の下で意味を与えることができる．

(c) 相関比

条件つき確率変数 $x_2|x_1$ の分散 $\sigma^2(x_2|x_1)$ は，R_2 の標本点の x_1 成分が与えられたとき，標本点の x_2 成分をうまく定める方法に関して重要な情報を与える．$\sigma^2(x_2|x_1)=0$ のときは確率 1 で定まる．$\sigma^2(x_2|x_1)\fallingdotseq 0$ のときは $\sigma^2(x_2)\fallingdotseq 0$ を仮定すると，比 $[\sigma^2(x_2|x_1)]/\sigma^2(x_2)$ をとることによって，うまく定められる．しかし，一般にこの比は x_1 に依存する．x_1 に関するこの比の平均値をとると，さらに良い基準が得られることがある．この平均値を $\eta_{2\cdot1}^2$ で表わせば，次のようになる．

(3.7.13)
$$\eta_{2\cdot1}^2=\frac{\mathscr{E}[\sigma^2(x_2|x_1)]}{\sigma^2(x_2)}.$$

ここで

(3.7.14)
$$\mathscr{E}[\sigma^2(x_2|x_1)]=\int_{-\infty}^{\infty}\sigma^2(x_2|x_1)\,dF_1(x_1).$$

$\eta_{2\cdot1}^2$ は x_1 上の x_2 の**相関比**と呼ばれる．$0\leqslant\eta_{2\cdot1}^2\leqslant1$ である．$\eta_{2\cdot1}^2=0$ となるのは $\sigma^2(x_2|x_1)=0$ のときに限る．すなわち，$P[(x_2=\mu(x_2|x_1))|x_1]=1$ のときのみである．一方 $\eta_{2\cdot1}^2=1$ となるのは x_1 と x_2 が無相関のときに限る．

重相関比 $\eta_{k\cdot12\cdots(k-1)}^2$ は次のように定義される．

(3.7.15)
$$\eta_{k\cdot12\cdots(k-1)}^2=\frac{\mathscr{E}[\sigma^2(x_k|x_1,\cdots,x_{k-1})]}{\sigma^2(x_k)}.$$

ここで

(3.7.16)
$$\mathscr{E}[\sigma^2(x_k|x_1,\cdots,x_{k-1})]$$
$$=\int_{R_{k-1}}\sigma^2(x_k|x_1,\cdots,x_{k-1})\,dF_{1\cdots(k-1)}(x_1,\cdots,x_{k-1}).$$

相関比 $\eta_{2\cdot1}^2$ の場合と同じく，$0\leqslant\eta_{k\cdot12\cdots(k-1)}^2\leqslant1$ となる．0 の値をとるのは $P[(x_k=\mu(x_k|x_1,\cdots,x_{k-1}))|x_1,\cdots,x_{k-1}]=1$ の場合のみで，値 1 をとるのは x_k と $(x_1,\cdots,x_{k-1})$ が無相関のときのみである．

3.8　最小 2 乗線形回帰

(a)　2 変数の場合

3.7.1 で述べた関数 $u(x_1)$ について，$\beta_0+\beta_1x_1$ の線形関数だけを考え，$\mathscr{E}(x_2-\beta_0-\beta_1x_1)^2$ が最小となる β_0 と β_1 を決定しよう．通常，x_1 も x_2 も退化しない，すなわち σ_1^2 および σ_2^2 はともに正であると仮定する．

この関数を $\phi(\beta_0,\beta_1)$ で表わせば，次のように書いてもよい．

(3.8.1)　$$\phi(\beta_0,\beta_1)=\mathscr{E}[(x_2-\mu_2)-(\beta_0-\mu_2+\beta_1\mu_1)-\beta_1(x_1-\mu_1)]^2.$$

ただし $\mu_1=\mathscr{E}(x_1)$ または $\mu_2=\mathscr{E}(x_2)$ である．$\phi(\beta_0,\beta_1)$ を最小にする β_0 と β_1 の値は次の等式で与えられよう．

(3.8.2)　$$\frac{\partial\phi}{\partial\beta_0}=0,\qquad \frac{\partial\phi}{\partial\beta_1}=0.$$

3.4 節の表現を用いて，これらの等式を簡単にすれば，次式が得られる．

(3.8.3)　$$\beta_0-\mu_2+\beta_1\mu_1=0,\qquad \rho\sigma_1\sigma_2-\beta_1\sigma_1^2=0.$$

その解を β_0^*，β_1^* で表わせば

(3.8.4)　$$\beta_0^*=\mu_2-\rho\frac{\sigma_2}{\sigma_1}\mu_1,\qquad \beta_1^*=\rho\frac{\sigma_2}{\sigma_1}.$$

したがって

3.8.1　$\mathscr{E}(x_2-\mu(x_1))^2$ を最小にする線形関数 $u(x_1)$ は次で与えられる．

(3.8.5)　$$u(x_1)=\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1).$$

直線の方程式は

(3.8.6)　$$x_2=\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1).$$

または，対称型に書いて

(3.8.7)　$$\frac{(x_2-\mu_2)}{\sigma_2}=\rho\frac{(x_1-\mu_1)}{\sigma_1}.$$

これは x_2 の x_1 上への**最小 2 乗回帰直線**と呼ばれる．

次のことが証明できる．

3.8.2　(3.8.4) で与えられる β_0 と β_1 の値で求まる $\mathscr{E}(x_2-\beta_0-\beta_1x_1)^2$ の最小値 $\sigma_{2\cdot1}^2$ は

(3.8.8)　$$\sigma_{2\cdot1}^2=\sigma_2^2(1-\rho^2)$$

で与えられる．

$\sigma_{2\cdot1}^2$ を x_2 の x_1 上への**最小 2 乗残差分散**と呼ぶ．それが 0 となるのは，$\rho=\pm1$ のとき，かつそのときに限る．すなわち x_1 と x_2 が（共に退化しないものと仮定して）**3.4.3** で示したように，真に 1 次従属であるときかつそのときに限る．

x_1 と x_2 を交換すれば，次の x_1 の x_2 上への**最小 2 乗回帰直線**が得られる．

(3.8.9)　$$\frac{(x_1-\mu_1)}{\sigma_1}=\rho\frac{(x_2-\mu_2)}{\sigma_2}.$$

x_1 も x_2 も退化しないから，2 つの回帰直線 (3.8.7) と (3.8.9) は，$\rho=\pm1$ のときかつそのときに限って一致することは明らかである．

ここで，その最小 2 乗回帰関数が，実際の回帰関数と一致する性質を持つ特殊な 2 次元の確率分布があるかどうかという疑問が起ってくる．この答は肯定的である．そして，数理統計学において，この性質を持つ最も重要な分布の 1 つは，**2 次元正規分布**すなわち**ガウス分布**であり，それは 7.3 節で説明する．この分布は，$\sigma^2(x_2|x_1)$ が x_1 によらず，実際には $\sigma_{2\cdot1}^2$ に等しいという性質を持つことがわかるだろう．

x_1 が与えられたときのある事象の x_2 の値を定めるために，最小 2 乗回帰直線がいかに効果的に用いられるかを示す基準は，x_2 の x_1 上への最小 2 乗回帰関数のまわりの x_2 の分散と，x_1 を無視した x_2 の分散とを比較することである．すなわち $\sigma_{2\cdot1}^2$ を σ_2^2 と比較することである．これら 2 つの量の比を次のように書いて，**1 次相関比**と呼ぶ．

(3.8.10)　$$\eta_{2\cdot1(L)}^2=\frac{\sigma_{2\cdot1}^2}{\sigma_2^2}=1-\rho^2.$$

x_1 と x_2 に線形な関係がある場合，すなわち x_1 と x_2 が真に 1 次従属な確率変数であれば，$\rho^2=1$ であり，$\eta_{2\cdot1(L)}^2=0$ である．逆に $\rho^2=1$（したがって $\eta_{2\cdot1(L)}^2=0$）ならば，x_1 と x_2 は 1 次従属である．一方，x_1 と x_2 の相関がなければ，$\rho=0$ で $\eta_{2\cdot1(L)}^2=1$ である．そして x_1 は x_2 を定める情報を含んでいない．

(b)　数個の変数の場合

さて次のような型の線形関数を考えてみよう．

(3.8.11)　$$\beta_0+\beta_1x_1+\cdots+\beta_{k-1}x_{k-1}.$$

いま，次式を最小にするように $\beta_0,\cdots,\beta_{k-1}$ を定めよう．

(3. 8. 12)　$$\phi(\beta_0, \beta_1, \cdots, \beta_{k-1}) = \mathscr{E}[x_k - \beta_0 - \beta_1 x_1 - \cdots - \beta_{k-1}x_{k-1}]^2.$$

2 次元の場合と同様に，(3. 8. 12) を次のように書く．

(3. 8. 13)　$$\phi = \mathscr{E}[(x_k - \mu_k) - (\beta_0 - \mu_k + \beta_1\mu_1 + \cdots + \beta_{k-1}\mu_{k-1}) - \beta_1(x_1 - \mu_1) - \cdots - \beta_{k-1}(x_{k-1} - \mu_{k-1})]^2.$$

(3. 8. 13) を最小にする β の値は次の等式の解で与えられる．

(3. 8. 14)　$$\frac{\partial\phi}{\partial\beta_\alpha} = 0, \qquad \alpha = 0, 1, \cdots, k-1.$$

$\partial\phi/\partial\beta_0 = 0$ は，次のようになる．

(3. 8. 15)　$$\beta_0 - \mu_k + \beta_1\mu_1 + \cdots + \beta_{k-1}\mu_{k-1} = 0.$$

(3. 8. 14) の残りの $k-1$ 個の式にこの結果を用いれば，次のようになる．

(3. 8. 16)　$$\begin{gathered} \beta_1\sigma_{11} + \cdots + \beta_{k-1}\sigma_{1k-1} = \sigma_{1k} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \beta_1\sigma_{k-1,1} + \cdots + \beta_{k-1}\sigma_{k-1,k-1} = \sigma_{k-1,k}. \end{gathered}$$

ここで $\|\sigma_{ij}\|$, $i, j = 1, \cdots, k$ は，3. 5 節で定義した確率変数 $(x_1, \cdots, x_k)$ の成分の共分散行列である．したがって

3. 8. 3　$|\sigma_{pq}| \neq 0$, $p, q = 1, \cdots, k-1$ のとき，ϕ を最小にする $\beta_0, \beta_1, \cdots, \beta_{k-1}$ の値を $\beta_0^*, \beta_1^*, \cdots, \beta_{k-1}^*$ で表わせば，それらは次式で与えられる．

(3. 8. 17)　$$\begin{gathered} \beta_0^* = \mu_k - \beta_1^*\mu_1 - \cdots - \beta_{k-1}^*\mu_{k-1} \\ \beta_p^* = \sum_{q=1}^{k-1} \sigma_{(k)}^{pq}\sigma_{qk}, \qquad p = 1, \cdots, k-1. \end{gathered}$$

ここで $\|\sigma_{(k)}^{pq}\|$ は共分散行列 $\|\sigma_{pq}\|$, $p, q = 1, \cdots, k-1$ の逆行列である．

3. 5. 1 から $|\sigma_{pq}| \neq 0$ は，$x_1, \cdots, x_{k-1}$ が 1 次従属でないことを示す．

最後に，x_k の $x_1, \cdots, x_{k-1}$ 上への最小 2 乗線形回帰超平面は，次のような式を持つ．

(3. 8. 18)　$$x_k = \mu_k + \sum_{p=1}^{k-1} \beta_p^*(x_p - \mu_p).$$

等式 (3. 8. 18) を，もっと簡単に行列式で表わすと次のようになる．

(3. 8. 19)　$$\begin{vmatrix} \sigma_{11} & \cdots & \sigma_{1k} \\ \vdots & & \vdots \\ \sigma_{k-1,1} & \cdots & \sigma_{k-1,k} \\ (x_1 - \mu_1) & \cdots & (x_k - \mu_k) \end{vmatrix} = 0$$

1 番下の行でこの行列式を展開すれば良くわかる．

さて，この β に関する $\phi(\beta_0, \cdots, \beta_{k-1})$ の最小値を求める問題を考えてみる．(3. 8. 17) から得られた β^* の値を (3. 8. 13) に代入して，ϕ の最小値を $\sigma_{k\cdot12\cdots(k-1)}^2$ で表わせば

(3. 8. 20)　$$\begin{aligned} \sigma_{k\cdot12\cdots(k-1)}^2 &= \mathscr{E}[(x_k - \mu_k) - \beta_1^*(x_1 - \mu_1) - \cdots - \beta_{k-1}^*(x_{k-1} - \mu_{k-1})]^2 \\ &= \sigma_{kk} - 2\sum_{p=1}^{k-1}\sigma_{pk}\beta_p^* + \sum_{p,q=1}^{k-1}\sigma_{pq}\beta_p^*\beta_q^*. \end{aligned}$$

(3. 8. 17) から得られる β_p^* の値を (3. 8. 20) の最後の 2 つの項に代入すれば，次の式が得られる．

(3. 8. 21)　$$\sum_{p=1}^{k-1}\sigma_{pk}\beta_p^* = \sum_{p,q=1}^{k-1}\sigma_{(k)}^{pq}\sigma_{pk}\sigma_{qk}$$

(3. 8. 22)　$$\sum_{p,q=1}^{k-1}\sigma_{pq}\beta_p^*\beta_q^* = \sum_{p,q=1}^{k-1}\sum_{r,s=1}^{k-1}\sigma_{pq}\sigma_{(k)}^{ps}\sigma_{(k)}^{qr}\sigma_{rk}\sigma_{sk} = \sum_{r,s=1}^{k-1}\sigma_{(k)}^{rs}\sigma_{rk}\sigma_{sk}.$$

なぜなら $\sum_{p=1}^{k-1}\sigma_{pq}\sigma_{(k)}^{ps} = \delta_{qs}$ は，$q = s$ のとき値 1 をとり，$q \neq s$ のときは 0 となるからである．ただし δ_{qs} はクロネッカーのデルタである．(3. 8. 22) の右端の項は，もちろん (3. 8. 21) の右辺と同じである．したがって次式が得られる．

(3. 8. 23)　$$\sigma_{k\cdot12\cdots(k-1)}^2 = \sigma_{kk} - \sum_{p,q=1}^{k-1}\sigma_{(k)}^{pq}\sigma_{pk}\sigma_{qk}.$$

しかし

(3. 8. 24)　$$\sigma_{kk} - \sum_{p,q=1}^{k-1}\sigma_{(k)}^{pq}\sigma_{pk}\sigma_{qk} = \frac{|\sigma_{ij}|}{|\sigma_{pq}|}.$$

ここで $i, j = 1, \cdots, k$ また $p, q = 1, \cdots, k-1$ である．これは行列式 $|\sigma_{ij}|$ を k 行と k 列で展開してみればわかる．[たとえば，Bôcher (1907) を参照せよ]．最終的に次の結果が得られる．

3. 8. 4　$\sigma_{k\cdot12\cdots(k-1)}^2$ で表わされる $\phi(\beta_0, \beta_1, \cdots, \beta_{k-1})$ の最小値は次のようになる．

(3. 8. 25)　$$\sigma_{k\cdot12\cdots(k-1)}^2 = \frac{|\sigma_{ij}|}{|\sigma_{pq}|}.$$

$\sigma_{k\cdot12\cdots(k-1)}^2$ は $x_1, \cdots, x_{k-1}$ 上の x_k の最小 2 乗残差分散と呼ばれる．それは式 (3. 8. 18) で表わされる最小 2 乗回帰平面のまわりの x_k の分散である．

確率変数 x_k と最小 2 乗回帰関数 $\mu_k + \sum_{p=1}^{k-1}\beta_p^*(x_p - \mu_p)$ との間の相関係数を x_k と

$(x_1, \cdots, x_{k-1})$ との間の重相関係数と呼ぶ．それを $\rho_{k\cdot 12\cdots(k-1)}$ で表わして $\|\sigma_{ij}\|$ の要素を用いれば次のように書ける．

$$\rho^2_{k\cdot 12\cdots(k-1)} = 1 - \frac{|\sigma_{ij}|}{\sigma_{kk}|\sigma_{pq}|}. \tag{3.8.26}$$

上の式を確かめるには，x_k の分散と $\mu_k + \sum_{p=1}^{k-1} \beta_p^*(x_p - \mu_p)$ の分散，およびそれらの間の共分散を定めよう．これは次のようになる．

$$\sigma^2(x_k) = \sigma_{kk}$$

$$\sigma^2\left[\mu_k + \sum_{p=1}^{k-1} \beta_p^*(x_p - \mu_p)\right] = \sum_{p,q=1}^{k-1} \sigma_{pq}\beta_p^*\beta_q^*$$

$$\mathrm{cov}\left[x_k, \mu_k + \sum_{p=1}^{k-1} \beta_p^*(x_p - \mu_p)\right] = \sum_{p=1}^{k-1} \sigma_{pk}\beta_p^*.$$

相関係数の定義 (3.4.6) を用いて，(3.8.21) と (3.8.22) および正の平方根を使えば，次が得られる．

$$\rho_{k\cdot 12\cdots(k-1)} = +\sqrt{\sum_{p,q=1}^{k-1} \sigma_{(k)}^{pq}\sigma_{pk}\sigma_{qk}/\sigma_{kk}}. \tag{3.8.27}$$

(3.8.27) に (3.8.24) を用いれば，式 (3.8.26) が得られる．

(3.8.25) と (3.8.26) から，明らかに次のことがいえる．

$$\sigma^2_{k\cdot 12\cdots(k-1)} = \sigma_{kk}(1 - \rho^2_{k\cdot 12\cdots(k-1)}). \tag{3.8.28}$$

1 次相関比は次のように定義される．

$$\eta^2_{k\cdot 12\cdots(k-1),(L)} = \frac{\sigma^2_{k\cdot 12\cdots(k-1)}}{\sigma_{kk}} = 1 - \rho^2_{k\cdot 12\cdots(k-1)}. \tag{3.8.29}$$

2 つの確率変数に対する 1 次相関比の場合と同様に，$\sigma^2_{k\cdot 12\cdots(k-1)}$ は $\rho^2_{k\cdot 12\cdots(k-1)} = 1$ のとき，かつそのときに限って 0 となる．すなわち $(x_1, \cdots, x_{k-1})$ 上の x_k の回帰平面に含まれる確率が 1 のときかつそのときに限る．

問　題

3.1 x が確率変数で，その 1 次の絶対中心モーメント ν_1 が存在すれば，x の平均 μ は有限であり，また $\lambda > 0$ に対して次式が成り立つことを示せ．

$$P(|x-\mu| < \lambda\nu_1) \geqq 1 - \frac{1}{\lambda}.$$

3.2 $g(x)$ が，確率変数 x の可測関数のとき，$\lambda > 0$ に対して $P(|g(x)| \geqq \lambda) \leqq$

$\mathscr{E}(|g(x)|)/\lambda$ となることを示せ．

3.3 確率変数 x の最初の r 個のモーメント $\mu_1', \cdots, \mu_r'$ (および中心モーメント $\mu_1, \cdots, \mu_r$) が存在するとき，次の関数式が成り立つことを示せ．

$$\mu_r' = \sum_{i=0}^{r}\binom{r}{i}\mu_{r-i}\mu^i, \qquad \mu_r = \sum_{i=0}^{r}(-1)^i\binom{r}{i}\mu_{r-i}'\mu^i.$$

3.4 x を質点 $0, 1, 2, \cdots, r$ を持つ任意の離散型確率変数とすれば，$\mu_{[r]}'$ より高次の階乗モーメントはすべて 0 であることを証明せよ．

3.5 確率変数のモーメント μ_{2r}，μ_{2r+1}，μ_{2r+2} が存在すれば，次式が成り立つことを示せ．

$$(\mu_{2r+1})^2 \leqq \mu_{2r}\mu_{2r+2}.$$

3.6 **3.5.2** および **3.5.3** を証明せよ．

3.7 $(x_1, \cdots, x_k)$ が k 次元確率変数で，平均 $(\mu_1, \cdots, \mu_k)$ および (正定値) 共分散行列 $\|\sigma_{ij}\|$ と逆行列 $\|\sigma^{ij}\|$ を持つとすれば，次式が成り立つことを証明せよ．

$$P\left(\sum_{i,j=1}^{k}\sigma^{ij}(x_i - \mu_i)(x_j - \mu_j) < \lambda^2\right) \geqq 1 - \frac{k}{\lambda^2}.$$

3.8 **3.6.1** を証明せよ．

3.9 $(x_1, \cdots, x_k)$ が独立な確率変数で，平均が 0 で，分散が 1 のとき，次のことを示せ．

$$P\left(\sum_{i=1}^{k} x_i^2 \geqq \lambda k\right) \leqq \frac{1}{\lambda}.$$

3.10 x は平均 μ と分散 σ^2 を持つ確率変数で，c.d.f. $F(x)$ を持てば，次のことが成り立つことを示せ．

$$F(x)\begin{cases} \leqq \dfrac{1}{1+\left(\dfrac{x-\mu}{\sigma}\right)^2}, & x < \mu \text{ のとき} \\[2ex] \geqq \dfrac{1}{1+\left(\dfrac{\sigma}{x-\mu}\right)^2}, & x > \mu \text{ のとき} \end{cases}$$

[Cramér (1946) を見よ].

3.11 2 つの確率変数 x と y の $2r$ 次のモーメントが有限ならば，$x+y$ の $2r$ 次のモーメントもまた有限であることを証明せよ．

3.12 x は平均 μ と分散 σ^2 を持つ確率変数で，かつ c で絶対最大値をとるような p.d.f. を持てば，次が成り立つことを証明せよ．

$$P(|x-c| \geqq \lambda\sqrt{\sigma^2 + (\mu - c)^2}) \leqq \left(\frac{2}{3\lambda}\right)^2.$$

この結果はガウスによる．

3.13 (続き) 次を証明せよ．

$$P(|x-\mu| \geqq \lambda\sigma) \leqq \frac{4(1+\delta^2)}{9[\lambda - |\delta|]^2}.$$

ただし $\delta = (\mu - c)/\sigma$, $\lambda > |\delta|$ である．[Cramér (1946)]．

3.14　$x_1, \cdots, x_n$ を，平均がすべて μ に等しく，分散がすべて σ^2 に等しい独立な確率変数とすると，任意の $\delta > 0$ に対して次が成り立つことをチェビシェフの不等式を用いて証明せよ．

$$\lim_{n \to \infty} P\left(\left|\frac{x_1 + \cdots + x_n}{n} - \mu\right| < \delta\right) = 1.$$

3.15　$(x_1, \cdots, x_k)$ が，おのおの 2 つの成分の間の相関係数が ρ である k 次元確率変数のとき，次を示せ．

$$-\frac{1}{k-1} \leqslant \rho \leqslant +1.$$

3.16　$(x_1, \cdots, x_m, y_1, \cdots, y_n)$ を $(m+n)$ 次元確率変数で，すべての成分の分散が 1, $\mathrm{cov}(x_i, x_j) = \rho_1$, $\mathrm{cov}(y_p, y_q) = \rho_2$, そして $\mathrm{cov}(x_i, y_q) = \rho_3$ とする．$u = x_1 + \cdots + x_m$, $v = y_1 + \cdots + y_n$ とすれば，u と v の間の相関係数は次のようになることを証明せよ．

$$\rho(u, v) = \frac{\sqrt{mn}\,\rho_3}{\sqrt{1 + (m-1)\rho_1} \cdot \sqrt{1 + (n-1)\rho_2}}.$$

3.17　$x_1, \cdots, x_p$, $y_1, \cdots, y_q$, $z_1, \cdots, z_r$ が分散 1 と共分散 0 を持つとき，$u = x_1 + \cdots + x_p + y_1 + \cdots + y_q$, $v = x_1 + \cdots + x_p + z_1 + \cdots + z_r$ とすれば，u と v の間の相関係数は次で与えられることを示せ．

$$\rho(u, v) = \frac{p}{\sqrt{(p+q)(p+r)}}.$$

3.18　$(x_1, \cdots, x_n)$ を，x_1 と条件つき確率変数 $x_2|x_1, \cdots, x_n|x_{n-1}$ とが独立であるような確率変数とする．このとき

$$\mathscr{E}(x_1) = \mu, \qquad \mathscr{E}(x_\xi | x_{\xi-1}) = x_{\xi-1}, \qquad \xi = 2, \cdots, n.$$

かつ
$$\mathscr{E}(x_1 - \mu)^2 = \mathscr{E}[(x_\xi - x_{\xi-1})^2 | x_{\xi-1}] = \sigma^2, \qquad \xi = 2, \cdots, n$$

ならば，x_n の条件のつかない平均および分散はそれぞれ μ および $n\sigma^2$ であることを示せ．

3.19　x を，最初の $2k$ 個のモーメント $\mu_1, \cdots, \mu_{2k}$ が存在する確率変数とすると，行列 $\|\mu_{i+j}\|$, $i, j = 1, \cdots, k$ は正定値であることを示せ．

3.20　x が非負の確率変数ならば，$\mathscr{E}(1/x) \geqslant 1/\mathscr{E}(x)$ となることを証明せよ．

3.21　$(x_1, \cdots, x_m)$ を（学年 G から "無作為" に選んだ）1 人の学生が，課目 A の試験での m 個の問題でとった得点を表わす m 次元確率変数とし，$(y_1, \cdots, y_n)$ を，課目 B の試験での得点を表わす n 次元確率変数とする．$T_1 = x_1 + \cdots + x_m$ および $T_2 = y_1 + \cdots + y_n$ をそれぞれ A および B における試験の総合 "得点" とする．$\sigma_{1,m}^2$ と $\sigma_{1,m}^2\rho_{1,m}$ はそれぞれ x の分散の平均と，すべての x の 2 組の共分散の平均を表わすものとする．$\sigma_{2,n}^2$ および $\sigma_{2,n}^2\rho_{2,n}$ を y に関する上記に相当する量とする．$\sigma_{1,m}\sigma_{2,n}\rho_{3,m,n}$ をすべての x, y の組の平均とする．$m, n \longrightarrow \infty$ のとき，$\sigma_{1,m}^2 \longrightarrow \sigma_1^2$, $\sigma_{2,n}^2 \longrightarrow \sigma_2^2$, $\rho_{1,m} \longrightarrow \rho_1$, $\rho_{2,m} \longrightarrow \rho_2$, $\rho_{3,m,n} \longrightarrow \rho_3$（ただし，$\sigma_1^2$, σ_2^2, ρ_1, ρ_2, ρ_3 はすべて正）ならば，次が成り立つことを示せ．

$$\lim_{m, n \to \infty} \rho(T_1, T_2) = \frac{\rho_3}{\sqrt{\rho_1\rho_2}}.$$

（したがって，T と T^* がその学生の同一課目（A または B）の 2 つの試験に関する長期にわたる得点とすれば，$\rho(T, T^*) \cong 1$.）

3.22　(x_1, x_2, x_3) は 3 次元確率変数で，（有限）共分散行列 $\|\sigma_{ij}\|$ と相関行列 $\|\rho_{ij}\|$, $i, j = 1, 2, 3$ を持つものとする．$\beta_{02}^* + \beta_{12}^* x_1$ を x_2 の x_1 上への最小 2 乗回帰直線，$\beta_{03}^* + \beta_{13}^* x_1$ を x_3 の x_1 上への最小 2 乗回帰直線とする．また $y_2 = x_2 - \beta_{02}^* - \beta_{12}^* x_1$, $y_3 = x_3 - \beta_{03}^* - \beta_{13}^* x_1$ とする．このとき，y_2 と y_3 の間の相関係数を $\rho_{23\cdot1}$ で表わし，x_1 を定数とした x_2 と x_3 の間の**偏相関係数**と呼ぶ．これに関して次のことを証明せよ．

$$\rho_{23\cdot1} = \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{(1 - \rho_{13}^2)(1 - \rho_{12}^2)}}.$$

3.23　上で述べた問題を一般化して，$(x_1, \cdots, x_k, x_{k+1})$ を $(k+1)$ 次元確率変数で，有限共分散行列 $\|\sigma_{ij}\|$ および相関行列 $\|\rho_{ij}\|$, $i, j = 1, \cdots, k$ を持つものとする．

$\beta_{0k}^* + \beta_{1k}^* x_1 + \cdots + \beta_{k-1,k}^* x_{k-1}$ を x_k の，さらに $\beta_{0\,k+1}^* + \beta_{1\,k+1}^* x_1 + \cdots + \beta_{k-1\,k+1}^* x_{k-1}$ を x_{k+1} の $x_1, \cdots, x_{k-1}$ 上への最小 2 乗回帰 "平面" とする．さらに

$$y_1 = x_k - \beta_{0k}^* - \beta_{1k}^* x_1 - \cdots - \beta_{k-1\,k}^* x_{k-1}$$
$$y_2 = x_{k+1} - \beta_{0\,k+1}^* - \beta_{1\,k+1}^* x_1 - \cdots - \beta_{k-1\,k+1}^* x_{k-1}$$

とする．$x_1, \cdots, x_{k-1}$ を定数とした x_k と x_{k+1} の間の偏相関係数 $\rho_{k,k+1\cdot12\cdots(k-1)}$ は y_1 と y_2 の間の相関係数である．このとき次を証明せよ．

$$\rho_{k,k+1\cdot12\cdots(k-1)} = \frac{\Delta_{k\,k+1}}{\sqrt{\Delta_{kk}\Delta_{k+1\,k+1}}}.$$

ただし，Δ_{pq}, $p, q = k, k+1$ は次に示す行列式の p 行および q 列を除去して得られる小行列式である．

$$\begin{vmatrix} 1 & \rho_{12} & \cdots & & \rho_{1,k+1} \\ \rho_{21} & 1 & & & \rho_{2,k+1} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 1 & \rho_{k,k+1} \\ \rho_{k+1,1} & \rho_{k+1,2} & \cdots & \rho_{k+1,k} & 1 \end{vmatrix}$$

3.24　$(x_1, \cdots, x_k)$ が k 次元確率変数で，おのおのの 2 組の成分の間の相関係数が ρ ならば，任意の成分とその残りの成分との間の重相関係数は次のようになることを証明せよ．

$$\rho \frac{\sqrt{k-1}}{\sqrt{1 + (k-2)\rho}}.$$

3.25　前の問題で，x_1 と成分 $x_2, \cdots, x_k$ のおのおのとの間の相関係数が ρ_1 で，成分 $(x_2, \cdots, x_k)$ のおのおのの 2 組の間の相関係数は ρ であるとする．このとき，$\rho_1\rho \geqslant 0$ を

示せ．また x_1 と $x_2, \cdots, x_k$ との間の重相関係数は次のようになることを示せ．

$$\frac{\sqrt{(k-1)\rho_1^2}}{\sqrt{1+(k-2)\rho}}.$$

3.26　問題 3.25 において，x_1 と x_r との間の相関係数を ρ_r，$r=2,\cdots,k$ とし，一方 $x_2, \cdots, x_k$ は互いに独立とする．このとき x_1 と $x_2, \cdots, x_k$ との間の重相関係数は $\sqrt{\rho_2^2+\cdots+\rho_k^2}$ となることを証明せよ．

3.27　問題 3.25 および 3.26 において，x_1 の $x_2, \cdots, x_k$ 上への最小 2 乗回帰超平面の式を最も簡単な形で示せ．

3.28　(3.8.18) と (3.8.19) によって与えられた式は同値であることを証明せよ．

第4章　確率変数列

4.1　確率過程の定義

数理統計学における重要な問題の 1 つに，n 個の確率変数の関数で，$n \longrightarrow \infty$ としたときの極限分布関数の決定がある．つまり $(x_1, \cdots, x_n)$ が n 次元確率変数であり，$g_n(x_1, \cdots, x_n)$ それ自身も確率変数である $(x_1, \cdots, x_n)$ の関数の場合，$n \longrightarrow \infty$ としたとき $g_n(x_1, \cdots, x_n)$ の c.d.f. の極限を決定すること，または少なくとも極限が存在するならば，この c.d.f. の性質を調べることが問題である．このように無限に多くの成分を持つ確率変数を扱うことができれば便利である．この確率変数を**確率過程**と呼び，可付番無限個の成分を持つ確率過程は**確率変数列**とみなせる．

いいかえれば，確率過程は確率変数の族 $\{x_\alpha ; \alpha \in A\}$ である．ただし，領域 A は，実数軸上の区間か，整数列のような実数軸上の点列であるか，あるいは成分 $(x_{\alpha_1}, \cdots, x_{\alpha_m})$ のどの有限の組合せも特定の c.d.f. $F_{\alpha_1, \cdots, \alpha_m}(x_{\alpha_1}, \cdots, x_{\alpha_m})$ を持つ確率変数の集合であるような，一般的な点集合である．確率変数のすべての有限個の組合せの c.d.f. は無矛盾でなければならない．すなわち，確率変数のどのような有限集合 —— たとえば $(x_{\alpha_1}, \cdots, x_{\alpha_m})$ —— の c.d.f. と，$(x_{\alpha_1}, \cdots, x_{\alpha_m})$ を部分集合とする確率変数の任意の有限集合の c.d.f. から得られたこの集合の周辺 c.d.f. とが同じ c.d.f. でなければならないという意味である．

4.2　確率過程に対する確率測度

ここでは成分の数が可付番無限個の確率過程 $(x_1, x_2, \cdots)$ を主に扱う．また，積空間 $R_1^{(1)} \times R_1^{(2)} \times \cdots$ を R_∞ で表わす．ただし $R_1^{(1)}, R_1^{(2)}, \cdots$ はおのおの $x_1, x_2, \cdots$ の（1次

元）標本空間である．すると，任意の実数の可付番無限個の数列 $(b_1, b_2, \cdots)$ も R_∞ における点であり，逆に R_∞ の任意の点は実数列である．このような点を b で表わす．それゆえ，可付番無限個の成分の値により求められた確率過程 $(x_1, x_2, \cdots)$ の実現値は R_∞ における標本点である．R_∞ はこのような確率過程の標本空間といえる．

b の任意の有限集合の座標 $(b_{\alpha_1}, \cdots, b_{\alpha_m})$ を考えてみる．この有限集合は，m 次元ユークリッド空間 $R_m^{(\alpha_1,\cdots,\alpha_m)} = R_1^{(\alpha_1)} \times \cdots \times R_1^{(\alpha_m)}$ の点 b' を表わす．点 b' は R_∞ の b から $R_m^{(\alpha_1,\cdots,\alpha_m)}$ への射影である．これを次のように書く．

(4.2.1)
$$b' = h_{\alpha_1,\cdots,\alpha_m}(b).$$

$R_m^{(\alpha_1,\cdots,\alpha_m)}$ における与えられた集合 E' へ射影する R_∞ の中のすべての点の集合 E は**筒集合**と呼ばれ，次のように書く．

(4.2.2)
$$E = h^{-1}_{\alpha_1,\cdots,\alpha_m}(E').$$

E' が $R_m^{(\alpha_1,\cdots,\alpha_m)}$ のボレル集合であれば，E' に対応する R_∞ での筒集合 E（R_∞ の E' の原像）は**ボレル筒集合**と呼ばれる．E' は単に R_∞ における E の $R_m^{(\alpha_1,\cdots,\alpha_m)}$ への射影である．ボレル筒集合の定義を $m = \infty$ の場合に拡張すれば，R_∞ と表わされ，それ自身もボレル筒集合になる．さらに，有限な m に対して，E がボレル筒集合であれば $\overline{E}$ もボレル筒集合になる．

ここで E', E'' を $R_m^{(\alpha_1,\cdots,\alpha_m)}$, $R_n^{(\beta_1,\cdots,\beta_n)}$ でのボレル集合とし，E_1, E_2 を R_∞ での対応する筒集合とする．R_∞ での集合 $E_1 \cup E_2$ について考える．これは R_∞ における $R_s^{(\gamma_1,\cdots,\gamma_s)}$ の $E_* = E'_* \cup E''_*$ に対応する筒集合である．ただし E'_* は $R_m^{(\alpha_1,\cdots,\alpha_m)}$ での E' に対応する $R_s^{(\gamma_1,\cdots,\gamma_s)}$ の筒集合であり，E''_* は $R_n^{(\beta_1,\cdots,\beta_n)}$ での E'' に対応する $R_s^{(\gamma_1,\cdots,\gamma_s)}$ の筒集合である．ただし，$\gamma_1,\cdots,\gamma_s$ は整数 $\alpha_1,\cdots,\alpha_m$, $\beta_1,\cdots,\beta_n$ を含む異なった整数の集合である．集合 $E_1 \cap E_2$ は $R_s^{(\gamma_1,\cdots,\gamma_s)}$ の $E'_* \cap E''_*$ に対応する R_∞ における筒集合である．E_1 と E_2 が互いに素であるのは $E'_* \cap E''_* = \varnothing$ の場合かつそのときのみである．

E', E'' がそれぞれ $R_m^{(\alpha_1,\cdots,\alpha_m)}$, $R_n^{(\beta_1,\cdots,\beta_n)}$ でのボレル集合で，その $R_s^{(\gamma_1,\cdots,\gamma_s)}$ の中の原像が E'_* と E''_* であれば，$E'_* \cup E''_*$ もまた $R_s^{(\gamma_1,\cdots,\gamma_s)}$ におけるボレル集合である．しかも，$E_1 \cup E_2$ は対応する R_∞ の筒集合であるので，この後者の集合はボレル筒集合である．ここで

4.2.1　R_∞ におけるボレル筒集合の族は集合のブール集合体 $\mathscr{F}$ である．

しかし，この集合体のボレル集合体への拡張（$\mathscr{B}_\infty$ と呼ぶ）が重要である．すなわち，ボレル集合体 $\mathscr{B}$ はブール集合体 $\mathscr{F}$ を含む R_∞ での集合の最小のボレル族として定義される．$\mathscr{B}_\infty$ は R_∞ の**ボレル集合族**と呼ばれ，R_∞ でのすべてのボレル筒集合を含んでいる．

ここで E' を空間 $R_m^{(\alpha_1,\cdots,\alpha_m)}$ でのボレル集合とし，かつ E を E' の R_∞ における原像とする．次のように，よく用いられる規則を適用してボレル筒集合 E に確率を定める．すなわち

(4.2.3)
$$P(E) = P_{\alpha_1,\cdots,\alpha_m}(E').$$

ただし $P_{\alpha_1,\cdots,\alpha_m}(E')$ は確率過程 $(x_1, x_2, \cdots)$ から取り出した成分 $(x_{\alpha_1}, \cdots, x_{\alpha_m})$ の c.d.f. から計算された E' に関する確率である．ボレル筒集合でない $\mathscr{B}_\infty$ の集合にどのような確率を対応させるべきかという問題が起って来る．この問題は Kolmogorov (1933 *a*) の次のような拡張定理が解決している．

4.2.2　$(x_{\alpha_1}, \cdots, x_{\alpha_m})$ を確率過程 $(x_1, x_2, \cdots)$ からの成分の任意の有限個の集まりとし，$R_m^{(\alpha_1,\cdots,\alpha_m)}$ を $(x_{\alpha_1}, \cdots, x_{\alpha_m})$ の標本空間とする．このとき，E が $R_m^{(\alpha_1,\cdots,\alpha_m)}$ での任意のボレル集合 E' に対応する R_∞ でのボレル筒集合であり，E に対する確率が (4.2.3) で与えられているならば，R_∞ のボレル集合族 $\mathscr{B}_\infty$ 上に確率測度が一意に存在する．この測度は R_∞ での筒集合については (4.2.3) で定義される集合関数である．

この定理の証明は省略する．証明に興味のある読者は Doob (1953) と Kolmogorov (1933 *a*) を参照せよ．

応用の場で数学モデルとして可付番無限個の成分を持つ確率過程を設定するとき，たとえば物理的過程の場合などでは，有限確率過程を持ち，4.2.2 の条件を満足するような方法で成分の列 $x_1, x_2, \cdots$ を“生成”することが考えられる．

例題　“ひずみのない”サイコロを何回も投げると簡単な確率過程 $(x_1, x_2, \cdots)$ が得られる．ここで x_α は α 回目に投げて出た目の数を表わす確率変数とする．$(x_{\alpha_1}, \cdots, x_{\alpha_r})$ は，任意の有限な r と任意に選んだ $\alpha_1, \cdots, \alpha_r$ に対して，r 次元確率変数 $(x_{\alpha_1}, \cdots, x_{\alpha_r})$ の 6^r 個の各質点上の関数 $p(x_{\alpha_1}, \cdots, x_{\alpha_r}) = 1/6^r$ を p.f. として持っている．

x_1 を $(0, 1)$ 上で，x_2 を $(x_1, 1)$ 上で，そして x_α を $(x_{\alpha-1}, 1)$, $\alpha = 1, 2, \cdots$ 上でそれぞれ矩形分布を持つ確率変数とすれば，$(x_1, x_2, \cdots)$ は **1 次のマルコフ連鎖**と呼ばれ，確率過程のうち最も重要な例の 1 つである．任意の α に対して条件つき確率変数 $x_\alpha | x_{\alpha-1}$ の p.d.f. $1/(1 - x_{\alpha-1})$ は正確には条件つき確率変数 $x_\alpha | x_{\alpha-1}, x_{\beta_1}, \cdots, x_{\beta_r}$ の p.d.f. と

同一である．ただし $\beta_1, \cdots, \beta_r$ は整数 $1, 2, \cdots, \alpha-2$ の $r < \alpha-2$ なる任意の集合である．k 次のマルコフ連鎖*)は，任意の α に対して，条件つき確率変数 $x_\alpha | x_{\alpha-1}, \cdots, x_{\alpha-k}$ が条件つき確率変数 $x_\alpha | x_{\alpha-1}, \cdots, x_{\alpha-k}, x_{\beta_1}, \cdots, x_{\beta_r}$ と同じ分布を持つ確率過程 $(x_1, x_2, \cdots)$ である．ただし $\beta_1, \cdots, \beta_r$ は整数 $1, \cdots, \alpha-k-1$ のうち $r < \alpha-k$ なる任意の集合である．マルコフ連鎖の一般理論およびマルコフ過程については，Doob (1953)，Feller (1957)，Loève (1955) を参照されたい．

確率過程 $(x_1, x_2, \cdots)$ が存在していて，x の有限個の集まりの (ユークリッド) 空間でのボレル集合に関する確率について論ずる場合には，R_∞ でのこれら集合の原像を常に考えることができ，それゆえ R_∞ での集合とその確率について常に論じることができる．たとえば $|x_m - x_n| < c$ に対する $x_m x_n$ 平面での (ボレル) 集合のかわりに，$|x_m - x_n| < c$ に対する R_∞ での (ボレル筒) 集合について論じることができる．2 つの集合に定めた確率は定義により等しい．特定な条件下では，$(x_1, x_2, \cdots)$ からの x の有限の集まりに対応する特定の有限次元空間での事象を扱うよりも，R_∞ での原像の事象を扱う方が便利である．$(x_1, \cdots, x_n)$ の標本空間が n により変化するのに対して，R_∞ は $(x_1, x_2, \cdots)$ に関する事象が起る固定した標本空間である．しかし，R_∞ で考えるのは便利さだけでなく，次節のように，R_∞ だけしか定式化できないある種の問題や定理があるためである．

4.3　確率収束

(a)　確率収束するためのいくつかの基準

$(x, x_1, x_2, \cdots)$ を任意の $\varepsilon > 0$ に対して次式が成り立つ確率過程とする．

(4.3.1) $$\lim_{n\to\infty} P(|x_n - x| \geqslant \varepsilon) = 0.$$

このとき $(x_1, x_2, \cdots)$ は確率変数 x に**確率過程的に収束する**または**確率収束する**といい，次式のように簡単に表わす．

(4.3.1 a) $$p \lim_{n\to\infty} x_n = x.$$

x が $P(x = x_0) = 1$ となる退化する確率変数ならば，確率過程 $(x_1, x_2, \cdots)$ は定数 x_0 に確率収束する．

*) 原書では "Markov chain of order k" であるが，"k 重 (多重) マルコフ連鎖" ということもある．(訳注)

定数に収束する確率過程の例で，最も簡単でかつ重要なものは次に述べる**大数の弱法則**である．

4.3.1　$(x_1, x_2, \cdots)$ を平均 0 と分散 $\sigma_1^2, \sigma_2^2, \cdots$ を持つ独立な確率変数列とする．ここで $c_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$ かつ $\lim_{n\to\infty} c_n^2/n^2 = 0$ とし，$\bar{x}_n = (x_1 + \cdots + x_n)/n$ とする．このとき確率過程 $(\bar{x}_1, \bar{x}_2, \cdots)$ は 0 に確率収束する．

これを証明するには $\mathscr{E}(\bar{x}_n) = 0$ と $\mathscr{E}(\bar{x}_n^2) = c_n^2/n^2$ に注目する．ここで任意の $\varepsilon > 0$ に対してチェビシェフの不等式を用いると

$$P(|\bar{x}_n| < \varepsilon) \geqslant 1 - \frac{c_n^2}{n^2\varepsilon^2}$$

となる．ここで $n \longrightarrow \infty$ のとき，$c_n^2/n^2 \longrightarrow 0$ ならば $\lim_{n\to\infty} P(|\bar{x}_n| < \varepsilon) = 1$ となり結論を得る．

実際には，$(x_1, x_2, \cdots)$ の各成分の間の共分散が 0 であれば，**4.3.1** は成り立つ．

$(x_1, x_2, \cdots)$ の成分の c. d. f. が $F_1(x), F_2(x), \cdots$ であり，かつ c. d. f. $F(x)$ の任意の連続点で $\lim_{n\to\infty} F_n(x) = F(x)$ であれば，$(x_1, x_2, \cdots)$ は $F(x)$ に**法則収束する**という．

確率収束および法則収束に関するいくつかの定理を次に述べる．これらは後章でも使用される．

確率収束する 1 番簡単な基準は次のようにいえる．

4.3.2　$(x, x_1, x_2, \cdots)$ が確率過程であり，かつ

(4.3.2) $$\lim_{n\to\infty} \mathscr{E}(x_n - x)^2 = 0$$

であれば，$(x_1, x_2, \cdots)$ は確率変数 x に確率収束する．

これは **3.3.2** で述べたチェビシェフの不等式より

(4.3.3) $$P(|x_n - x| \geqslant \varepsilon) \leqslant \frac{\mathscr{E}(x_n - x)^2}{\varepsilon^2}$$

となり，かつ (4.3.2) の仮定から (4.3.1)，すなわち，$\lim_{n\to\infty} P(|x_n - x| \geqslant \varepsilon) \leqslant 0$ が成り立つ．$(x, x_1, x_2, \cdots)$ が $\mathscr{E}(x_n^2) < +\infty$，$n = 1, 2, \cdots$ なる (4.3.2) と $\mathscr{E}(x^2) < +\infty$ を満足する確率過程であれば，$(x_1, x_2, \cdots)$ は x に**平均収束する**という．この形の収束を次のように表わすこともある

(4.3.2 a) $$\underset{n\to\infty}{\text{l. i. m.}}\, x_n = x.$$

4.3.3　$(x, x', x_1, x_2, \cdots)$ を $(x_1, x_2, \cdots)$ が確率変数 x と x' に確率収束する確率過程とすれば，x と x' は同値である．

この証明はまず

$$(4.3.4)\qquad P\left(|x-x'|>\frac{1}{N}\right)=P\left(|(x_n-x')-(x_n-x)|>\frac{1}{N}\right)$$

であり，E, E', E'' を $(x, x', x_1, x_2, \cdots)$ の標本空間 R_∞ でのそれぞれ，$|(x_n-x')-(x_n-x)|>\frac{1}{N}$，$|x_n-x|>\frac{1}{2N}$，$|x_n-x'|>\frac{1}{2N}$ が成り立っている点集合とする．もちろん，$E\subset(E'\cup E'')$ である．よって $P(E)\leqslant P(E')+P(E'')$ が成り立つ．すなわち

$$(4.3.5)\qquad P\left(|x-x'|>\frac{1}{N}\right)\leqslant P\left(|x_n-x|>\frac{1}{2N}\right)+P\left(|x_n-x'|>\frac{1}{2N}\right).$$

いま $n\longrightarrow\infty$ とすると，仮定により右辺の 2 つの項の極限は 0 になる．ゆえに

$$P\left(|x-x'|>\frac{1}{N}\right)=0.$$

$|x-x'|>\frac{1}{N}$ となる R_∞ での集合を G_N と表わすと，$G_1\subset G_2\subset\cdots$ かつ $\lim_{N\to\infty}G_N=G$ が得られる．ただし G は $x\neq x'$ となる R_∞ の点集合である．それゆえ，$\mathscr{B}_\infty$ の集合の場合の **1.4.6** により

$$(4.3.6)\qquad P(x\neq x')=P(G_1\cup G_2\cup\cdots)\leqslant P(G_1)+P(G_2)+\cdots$$

を得る．しかし $P(G_1)=P(G_2)=\cdots=0$ なので $P(x\neq x')=0$，よって x と x' は同値である．[2.8 (b) 節を見よ]．

もう 1 つの重要な結果は確率収束が法則収束を含むことである．すなわち，

4.3.4　$(x, x_1, x_2, \cdots)$ は各成分がそれぞれ $F(x), F_1(x), F_2(x), \cdots$ なる c.d.f. を持つ確率過程とする．$(x_1, x_2, \cdots)$ が確率変数 x に確率収束するならば，c.d.f. の列 $F_1(x), F_2(x), \cdots$ は $F(x)$ の連続な各点において $F(x)$ に収束する．

R_∞ を確率過程 $(x, x_1, x_2, \cdots)$ の標本空間とする．$F(x)$ は $x=x_0$ で連続で，x' は $x'<x_0$ なる定数とする．ここで R_∞ において $x\leqslant x'$，$x_n\leqslant x_0$，$|x_n-x|>(x_0-x')$ なる 3 つの事象を考える．最初の事象は第 2 の事象および第 3 の事象の和事象に含まれる．それゆえ

$$(4.3.7)\qquad F(x')=P(x\leqslant x')\leqslant P(x_n\leqslant x_0)+P(|x_n-x|>(x_0-x')).$$

ここで $n\longrightarrow\infty$ として極限をとる．$(x_1, x_2, \cdots)$ は x に確率収束するので $\lim_{n\to\infty}P(|x_n-x|>(x_0-x'))=0$ となる．それゆえ

$$(4.3.8)\qquad F(x')\leqslant\liminf_{n\to\infty}F_n(x_0).$$

同様に $x''>x_0$ とすると

$$(4.3.9)\qquad F(x'')\geqslant\limsup_{n\to\infty}F_n(x_0).$$

ゆえに

$$(4.3.10)\qquad F(x')\leqslant\liminf_{n\to\infty}F_n(x_0)\leqslant\limsup_{n\to\infty}F_n(x_0)\leqslant F(x'').$$

しかし $x=x_0$ で $F(x)$ は連続なので，$\lim_{x'\to x_0}F(x')=\lim_{x''\to x_0}F(x'')=F(x_0)$ となる．それゆえ

$$(4.3.11)\qquad \lim_{n\to\infty}F_n(x_0)=F(x_0).$$

これで **4.3.4** を得た．

(b)　確率過程における成分の関数の収束

$(x, x_1, x_2, \cdots)$ は，$(x_1, x_2, \cdots)$ が x に確率収束する確率過程とする．この場合，$(g(x_1), g(x_2), \cdots)$ が $g(x)$ へ確率収束するためにどのような条件が必要かを知る必要があるだろう．これに関するいくつかの定理と，今後用いる同様な問題を考えよう．

4.3.5　$(x, x_1, x_2, \cdots)$ は $(x_1, x_2, \cdots)$ が x へ確率収束する確率過程とする．$g(x)$ を R_1 上の x の連続関数とする．このとき

$$(4.3.12)\qquad p\lim_{n\to\infty}g(x_n)=g(x).$$

x が定数 c へ退化する確率変数ならば，(4.3.12) において $(g(x_1), g(x_2), \cdots)$ は定数 $g(c)$ へ確率収束する．

4.3.5 の証明には，$g(x)$ が任意の閉区間たとえば $[-M, M]$ 上で一様連続であることに注意する．x が確率変数であるとき，任意の $\varepsilon>0$ に対して適当に M を選べば

$$(4.3.13)\qquad P(|x|>M)<\frac{\varepsilon}{2}.$$

このように ε と M を選べば，次のような $\delta(\varepsilon, M)$ が存在する．すなわち

(i)　$|x|\leqslant M$

(ii)　$|x_n-x|<\delta(\varepsilon, M)$

x　x'　x_0　x_n

$|x_n-x$

$x_n-x>x_0-x'$　　$x-x_n>x_0-x'$

x_n-x

であれば

$$\text{(iii)}\quad |g(x_n)-g(x)|<\varepsilon.$$

(i), (ii), (iii) を満たす R_∞ での集合はそれぞれ E_1, E_2, E_3 で表わせば，$\overline{E_3}\subset(\overline{E_1}\cup\overline{E_2})$ となり，$P(\overline{E_3})\leqslant P(\overline{E_1})+P(\overline{E_2})$ を得る．すなわち

$$(4.3.14)\quad P(|g(x_n)-g(x)|\geqslant\varepsilon)\leqslant P(|x_n-x|\geqslant\delta(\varepsilon,M))+P(|x|>M).$$

しかし $n>n(\varepsilon,\delta,M)$ であれば，次式を満たすような $n(\varepsilon,\delta,M)$ が存在する．

$$(4.3.15)\quad P(|x_n-x|\geqslant\delta(\varepsilon,M))<\frac{\varepsilon}{2}.$$

M は (4.3.13) を満足するように選ばれているので，最終的には $n>n(\varepsilon,\delta,M)$ に対して

$$(4.3.16)\quad P(|g(x_n)-g(x)|\geqslant\varepsilon)<\varepsilon$$

を得る．これは (4.3.12) と同値である．よって **4.3.5** が得られた．

4.3.5 は $g(x)$ を $P(x\in I)=1$ なる閉区間 I 上で連続とすることにより，さらに一般化できる．これは論旨を多少変えれば得られるので，読者にまかせる．

成分がベクトルの確率過程の収束問題を取り扱わなくてはならないときがあるだろう．たとえば，$(x,y;x_1,y_1;x_2,y_2;\cdots)$ がこのような確率過程であるならば $(g(x_1,y_1), g(x_2,y_2),\cdots)$ が $g(x,y)$ に確率収束するのは，どういう条件下であるだろうか．この場合 R_∞ は $(x,y;x_1,y_1;x_2,y_2;\cdots)$ の標本空間である．

$(z;x_1,y_1;x_2,y_2;\cdots)$ が任意の $\varepsilon>0$ に対して

$$\lim_{n\to\infty}P(|x_n-y_n|>\varepsilon)=0$$

なる確率過程であり，列 $(x_1,x_2,\cdots)$ または $(y_1,y_2,\cdots)$ の一方が z に確率収束するならば，他方も z に確率収束する．このような場合，$(x_1,x_2,\cdots)$ と $(y_1,y_2,\cdots)$ とは<u>z に共に確率収束する</u>という．同じく，上の極限に関する式が成り立っていて，2 つの列のうちの一方が $F(x)$ へ法則収束するならば，2 つの列は<u>$F(x)$ へ共に法則収束する</u>という．

次に示すようにいくつかの確率変数列の場合にも **4.3.5** を拡張しよう．すなわち

4.3.6　$(x^{(1)},\cdots,x^{(k)};x_1^{(1)},\cdots,x_1^{(k)};x_2^{(1)},\cdots,x_2^{(k)};\cdots)$ は k 次元確率変数 $(x^{(1)},\cdots,x^{(k)})$ へ確率収束する k 次元ベクトル値確率過程とする．$g(x^{(1)},\cdots,x^{(k)})$ を R_k での $(x^{(1)},\cdots,x^{(k)})$ の連続関数とすれば，次式が成り立つ．

$$(4.3.17)\quad p\lim_{n\to\infty}g(x_n^{(1)},\cdots,x_n^{(k)})=g(x^{(1)},\cdots,x^{(k)}).$$

この定理の証明は **4.3.5** の証明と同じであるので，読者にまかす．

さらに **4.3.6** の条件を，$g(x^{(1)},\cdots,x^{(k)})$ が $P((x^{(1)},\cdots,x^{(k)})\in I_k)=1$ となる k 次元閉区間 I_k で連続である，とゆるめることができる．

次に示す **4.3.6** の系は，次節でよく利用される．

4.3.6 a　$(x,c;x_1,y_1;x_2,y_2;\cdots)$ はその $(x_1,y_1;x_2,y_2;\cdots)$ が (x,c) に確率収束するベクトル値確率過程であれば，$F(x)$ の連続な各点で

$$(4.3.18)\quad \lim_{n\to\infty}P(x_n+y_n\leqslant z)=F(z-c)$$

$$(4.3.19)\quad \lim_{n\to\infty}P(x_ny_n\leqslant z)=F\left(\frac{z}{c}\right)$$

$$(4.3.20)\quad \lim_{n\to\infty}P\left(\frac{x_n}{y_n}\leqslant z\right)=F(cz)$$

$$(4.3.21)\quad \lim_{n\to\infty}P(ax_n+by_n\leqslant z)=F\left(\frac{z-bc}{a}\right)$$

となる．ただし，x は c.d.f. $F(x)$ を持ち，かつ c は正の定数である．また，a と b は定数でかつ $a>0$.

次の定理も後章で利用される．

4.3.7　$(x_1^{(1)},\cdots,x_1^{(k)};y_1^{(1)},\cdots,y_1^{(k)};x_2^{(1)},\cdots,x_2^{(k)};y_2^{(1)},\cdots,y_2^{(k)};\cdots)$ はベクトル値確率過程で

$$(4.3.22)\quad |y_n^{(i)}-c^{(i)}|\leqslant|x_n^{(i)}-c^{(i)}|,\quad i=1,\cdots,k;\quad n=1,2,\cdots$$

を満足し，かつ $(x_1^{(1)},\cdots,x_1^{(k)};x_2^{(1)},\cdots,x_2^{(k)};\cdots)$ は定数ベクトル $(c^{(1)},\cdots,c^{(k)})$ に確率収束するとする．また $g(x^{(1)},\cdots,x^{(k)})$ が R_k でのすべての点で 1 価関数で，かつ $(c^{(1)},\cdots,c^{(k)})$ を含む k 次元開矩形区間で連続ならば

$$(4.3.23)\quad p\lim_{n\to\infty}g(y_n^{(1)},\cdots,y_n^{(k)})=g(c^{(1)},\cdots,c^{(k)}).$$

この定理の証明は容易なので省略する．**4.3.7** での条件は，$g(x^{(1)},\cdots,x^{(k)})$ を $P(x^{(1)},\cdots,x^{(k)}\in I_k)=1$ なる $(c^{(1)},\cdots,c^{(k)})$ を含む k 次元閉区間 I_k においてのみ定義され，かつ I_k において $g(x^{(1)},\cdots,x^{(k)})$ が連続であるという条件になおせる．

$(g(x,\theta),g(x_1,\theta),g(x_2,\theta),\cdots)$ はある区間 (θ',θ'') の各 θ に対して確率過程であるとする．任意の $\varepsilon>0$ に対して，1 つの n_ε があって，$n>n_\varepsilon$ ならば，区間 (θ',θ'') の任意の θ に対して

(4.3.24)　$$P(|g(x_n,\theta)-g(x,\theta)|>\varepsilon)<\varepsilon$$

を満足するとき，$(g(x_1,\theta), g(x_2,\theta),\cdots)$ は区間 (θ',θ'') の θ に関して $g(x,\theta)$ に一様に確率収束するという．このことは，$g(x,\theta)$ が退化する場合でも，また x_n がベクトル確率変数の場合にも，この方法を用いて拡張できる．

最後に，次の定理は後章で，特に第 12，13 章で用いる．

4.3.8　$(x_1,x_2,\cdots)$ をパラメータ θ に従属する確率過程とする．またある区間 (θ',θ'') における θ の値に対して，$f_n(x_1,\cdots,x_n,\theta)$，$n=1,2,\cdots$ と $\theta_n^*(x_1,\cdots,x_n)$，$n=1,2,\cdots$ は $g(\theta)$ と θ に一様に確率収束する確率変数列である．ただし $g(\theta)$ は (θ',θ'') で連続である．このとき，$f_n(x_1,\cdots,x_n,\theta_n^*)$，$n=1,2,\cdots$ は $g(\theta)$ へ確率収束する．

(θ',θ'') の任意の点 θ_0 における確率過程 $(x_1,x_2,\cdots)$ を考える．列 $(\theta_1^*(x_1),\theta_2^*(x_1,x_2),\cdots)$ は θ_0 に確率収束する．すなわち $(\theta_0-\varepsilon,\theta_0+\varepsilon)$ が，(θ',θ'') に含まれるような任意の $\varepsilon>0$ に対して $n>n_\varepsilon$ ならば，次式を満足する $(\theta_n^*(x_1,\cdots,x_n)$ を θ_n^* と書く) 1 つの n_ε が存在する．

(4.3.25)　$$P(\theta_0-\varepsilon<\theta_n^*<\theta_0+\varepsilon)>1-\varepsilon.$$

いま $(\theta_0-\varepsilon,\theta_0+\varepsilon)$ における θ の値に対する $g(\theta)$ の最小上界を $g_u(\theta_0)$，最大下界を $g_l(\theta_0)$ とする．$(f_1(x_1,\theta), f_2(x_1,x_2,\theta),\cdots)$ は (θ',θ'') の θ に関して $g(\theta)$ へ一様に確率収束するので，任意の $\varepsilon'>0$ に対して $n_{\varepsilon'}$ が存在し，$n>n_{\varepsilon'}$ ならば $(\theta_0-\varepsilon,\theta_0+\varepsilon)$ の任意の θ に対して次式が成立する．

(4.3.26)　$$P(g_l(\theta_0)-\varepsilon'<f_n(x_1,\cdots,x_n,\theta)<g_u(\theta_0)+\varepsilon')>1-\varepsilon'.$$

任意の $n>\max(n_\varepsilon,n_{\varepsilon'})$ に対して，E_n を $(x_1,x_2,\cdots)$ の標本空間において (4.3.25) と (4.3.26) の不等式をともに満足する R_∞ の集合とする．$P(E_n)>1-\varepsilon-\varepsilon'$ は明らかである．しかし E_n での任意の点は，次の不等式を満足する．

(4.3.27)　$$g_l(\theta_0)-\varepsilon'<f_n(x_1,\cdots,x_n,\theta_n^*)<g_u(\theta_0)+\varepsilon'.$$

ゆえに $n>\max(n_\varepsilon,n_{\varepsilon'})$ ならば，(4.3.27) が $(\theta_0-\varepsilon,\theta_0+\varepsilon)$ の任意の θ_n^* に対して成り立っている確率は $1-\varepsilon-\varepsilon'$ より大きい．しかし ε と ε' は任意であり，かつ $g(\theta)$ は (θ',θ'') で連続なので，差 $g_n(\theta_0)-g(\theta_0)$ および $g(\theta_0)-g_l(\theta_0)$ は任意に小さくできる．したがって，$f_n(x_1,\cdots,x_n,\theta^*)$ は $g(\theta_0)$ に確率収束する．しかも θ_0 は (θ',θ'') の中の任意の点である：これで **4.3.8** は証明された．

4.4　概　収　束

(a)　定　　義

$(x_1,x_2,\cdots)$ を確率空間 $(R_\infty,\mathscr{B}_\infty,P)$ に関する確率過程とし，E を収束する列 $(b_1,b_2,\cdots)$ のすべての集合とする．どの点が E に属するかを見るために，N が正の整数のとき，E_{npN} を

(4.4.1)　$$|b_{n+j}-b_n|<\frac{1}{N},\qquad j=1,2,\cdots,p$$

となる R_∞ の集合とする．

各 n，p，N に対して，E_{npN} は R_∞ でのボレル筒集合であり，もちろん $\mathscr{B}_\infty$ に属する．集合 $E_{(N)}=\bigcup_{n=1}^\infty\bigcap_{p=1}^\infty E_{npN}$ もまた $\mathscr{B}_\infty$ に属し，$N=1,2,\cdots$ に関して減少列である．上で述べた点集合 E は減少列の極限である．すなわち

(4.4.2)　$$E=\lim_N E_{(N)}$$

である．これもまた $\mathscr{B}_\infty$ に属し，確率 $P(E)$ を持つ．この確率は確率過程 $(x_1,x_2,\cdots)$ の収束確率と呼ばれる．E を R_∞ での収束集合という．

ここで，$P(E)=1$ で，x を $(b_1,b_2,\cdots)$ が E に属せば $\lim_{n\to\infty} b_n$ の値をとり，E に属さなければ値 0 をとる確率変数とする．このとき確率変数列は確率変数 x に概収束するという．

注意　前出の基本確率空間 $(R,\mathscr{B},P)$ での事象に関して概収束の定義を考えてみよう．一般に，R の標本点を e で表わすと確率過程（すべての成分が $\mathscr{B}$ に関して可測である）は $(x_1(e),x_2(e),\cdots)$ と書ける．すなわち，確率空間 $(R_\infty,\mathscr{B}_\infty,P)$ に関する確率過程に対して，R の各標本点 e を R_∞ の中の点へ写像する確率変数の可付番無限列が存在する．$(R_\infty$ での) 集合 E の (R での) 原像 E' の各点 e に対して $\lim_{n\to\infty} x_n(e)=x(e)$ となる確率変数 $x(e)$ ($\mathscr{B}$ に関して可測) が存在すれば，$(x_1(e),x_2(e),\cdots)$ は確率 $P(E')$ で $x(e)$ に収束するという．E' は R における収束集合である．$P(E')=P(E)=1$ であれば，$(x_1(e),x_2(e),\cdots)$ は $x(e)$ に概収束する，または確率 **1** で収束するという．

(b)　概収束と確率収束との関係

概収束は確率収束よりも強い収束である．すなわち

4.4.1　$(x,x_1,x_2,\cdots)$ が x に概収束する確率過程であれば，この列は x に確率収

束する．

なぜならば，E が $(x_1, x_2, \cdots)$ の収束集合であれば，任意の $\varepsilon > 0$ と n に対して，E を次に示す R_∞ での拡大集合列の極限として定義する．すなわち，$[|x_{n+j} - x| < \varepsilon,\ j = 0, 1, 2, \cdots]$，$n = 1, 2, \cdots$．しかし，$E_n$，$F_n$ をそれぞれ R_∞ の事象 $[|x_{n+j} - x| < \varepsilon,\ j = 0, 1, 2, \cdots]$，$[|x_n - x| < \varepsilon]$ で表わせば，$E_n \subset F_n$ である．よって

(4.4.3)
$$P(|x_n - x| < \varepsilon) \geqslant P(|x_{n+j} - x| < \varepsilon,\ j = 0, 1, 2, \cdots).$$

$n \longrightarrow \infty$ として極限をとると

(4.4.4)
$$\lim_{n\to\infty} P(|x_n - x| < \varepsilon) \geqslant \lim_{n\to\infty} P(|x_{n+j} - x| < \varepsilon,\ j = 0, 1, 2, \cdots) = P(\lim_{n\to\infty} E_n) = P(E) = 1$$

となる．すなわち

(4.4.5)
$$\lim_{n\to\infty} P(|x_n - x| < \varepsilon) = 1.$$

これは (4.3.1) と同値である．これで **4.4.1** は証明された．

4.5 コルモゴロフの不等式

Kolmogorov (1928) は次のようにチェビシェフの不等式の概念を複数個の不等式について拡張した．

4.5.1　$x_1, \cdots, x_n$ は平均 0，分散 $\sigma_1^2, \cdots, \sigma_n^2$ を持つ独立な確率変数の組とする．$c_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$ とすると，不等式

(4.5.1)
$$|x_1 + \cdots + x_\alpha| < \lambda c_n, \qquad \alpha = 1, \cdots, n$$

がすべて成り立つ確率は少なくとも $1 - 1/\lambda^2$ である．

証明は，F_α を次式を満たす標本空間 R_n での事象とする．

$$|x_1 + \cdots + x_\alpha| \geqslant \lambda c_n, \qquad \alpha = 1, \cdots, n.$$

$E_1 = F_1, E_2 = \overline{F}_1 \cap F_2, E_3 = \overline{F}_1 \cap \overline{F}_2 \cap F_3, \cdots, E_n = \overline{F}_1 \cap \overline{F}_2 \cap \cdots \cap \overline{F}_{n-1} \cap F_n$ とし，G_n を $E_1 \cup E_2 \cup \cdots \cup E_n$ の余集合とする．G_n は (4.5.1) の不等式のすべてが成り立つ R_n での集合である．$E_1, \cdots, E_n$，G_n は互いに素であり，それらの和集合は標本空間 R_n 全体である．すなわち

(4.5.2)
$$P(G_n) = 1 - [P(E_1) + \cdots + P(E_n)].$$

$\alpha = 1, \cdots, n$ に対して z_α を E_α における各点で値 1 をとり，$\overline{E}_\alpha$ の各点で値 0 をとる確率変数とし，z_{n+1} を G_n で同様に定義された確率変数とする．このとき

(4.5.3)
$$\mathscr{E}(x_1 + \cdots + x_n)^2 = \mathscr{E}[z_1(x_1 + \cdots + x_n)^2] + \cdots + \mathscr{E}[z_n(x_1 + \cdots + x_n)^2] + \mathscr{E}[z_{n+1}(x_1 + \cdots + x_n)^2].$$

$(x_1, \cdots, x_\alpha)$ の任意の関数 g_α は他の x に対しては独立なので

$$\mathscr{E}(g_\alpha \cdot x_\beta) = \mathscr{E}(g_\alpha) \cdot \mathscr{E}(x_\beta) = 0, \qquad \beta = \alpha + 1, \cdots, n.$$

それゆえ

(4.5.4)
$$\mathscr{E}[z_\alpha(x_1 + \cdots + x_n)^2] = \mathscr{E}[z_\alpha(x_1 + \cdots + x_\alpha)^2] + \mathscr{E}[z_\alpha(x_{\alpha+1} + \cdots + x_n)^2] \geqslant \mathscr{E}[z_\alpha(x_1 + \cdots + x_\alpha)^2].$$

しかし，E_α での各点では，$(x_1 + \cdots + x_\alpha)^2 \geqslant \lambda^2 c_n^2$．よって，$\alpha = 1, \cdots, n$ に対して

(4.5.5)
$$\mathscr{E}[z_\alpha(x_1 + \cdots + x_n)^2] \geqslant \lambda^2 c_n^2 P(E_\alpha).$$

$\mathscr{E}(x_1 + \cdots + x_n)^2 = c_n^2$ であるから，$\alpha = 1, \cdots, n$ に対して，(4.5.5) を (4.5.3) の右辺に代入して $\mathscr{E}[z_{n+1}(x_1 + \cdots + x_n)^2]$ を消去すると

(4.5.6)
$$c_n^2 \geqslant \lambda^2 c_n^2 [P(E_1) + \cdots + P(E_n)]$$

となる．(4.5.6) と (4.5.2) から

(4.5.7)
$$P(G_n) \geqslant 1 - \frac{1}{\lambda^2}.$$

これで **4.5.1** の証明は終わった．

4.6 大数の強法則

4.3 節で大数の弱法則について述べたが，これはある条件のもとで，平均 0 の独立な確率変数列の平均値確率変数が 0 に確率収束するということであった．これに対して，平均値列において，ある十分大きな番号以上の平均値のすべてが，任意に 0 に近づくことが確率的にいえるかどうかという問題が考えられる．これには次に示す大数の強法則が解を与えてくれる．

4.6.1　$(x_1, x_2, \cdots)$ を平均 0 と $\displaystyle\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < +\infty$ なる分散 $(\sigma_1^2, \sigma_2^2, \cdots)$ を持つ独立な

確率変数とする．また $\bar{x}_n=(x_1+\cdots+x_n)/n$ であれば，$(\bar{x}_1, \bar{x}_2, \cdots)$ は0に概収束する．すなわち任意の $\delta>0$ と $\varepsilon>0$ に対して $N_{\delta,\varepsilon}$ が存在し，すべての $N>N_{\delta,\varepsilon}$ と k に対して次式が成り立つ．

(4.6.1)　$$P(|\bar{x}_n|<\varepsilon, n=N, N+1, \cdots, N+k) \geqslant 1-\delta.$$

4.6.1 の証明は Feller (1957) と同じ方法で進める．$E_N, E_{N+1}, \cdots, E_{N+k}$ をそれぞれ $n=N, N+1, \cdots, N+k$ に対する $|\bar{x}_n|<\varepsilon$ なる事象とする．このとき (4.6.1) の確率は次式になる．

(4.6.2)　$$P(E_N \cap E_{N+1} \cap \cdots \cap E_{N+k}).$$

しかしこの確率は次の値を持つ．

(4.6.3)　$$1-P(\overline{E}_N \cup \overline{E}_{N+1} \cup \cdots \cup \overline{E}_{N+k}).$$

ここで $N>N_{\delta,\varepsilon}$，およびすべての k に対して，次式が成り立つことをいう．

(4.6.4)　$$P(\overline{E}_N \cup \overline{E}_{N+1} \cup \cdots \cup \overline{E}_{N+k}) \leqslant \delta.$$

I_α が $\{2^{\alpha-1}+1, 2^{\alpha-1}+2, \cdots, 2^\alpha\}$ となるように，正の整数全体を集合 $I_1, I_2, \cdots$ に分割する．F_α を不等式 $\{|\bar{x}_n|<\varepsilon,\ n \in I_\alpha\}$ の少なくとも1つが成り立たない事象とする．よってある α と β に対して

(4.6.5)　$$F_\alpha \cup F_{\alpha+1} \cup \cdots \cup F_{\alpha+\beta} \supset \overline{E}_N \cup \overline{E}_{N+1} \cup \cdots \cup \overline{E}_{N+k}.$$

また

(4.6.6)　$$P(F_\alpha \cup F_{\alpha+1} \cup \cdots \cup F_{\alpha+\beta}) \leqslant P(F_\alpha)+\cdots+P(F_{\alpha+\beta}).$$

これは $\sum_{\alpha=1}^{\infty} P(F_\alpha)$ が収束するのに十分である．しかも，F_α は I_α の少なくとも1つの n に対して $|\bar{x}_n| \geqslant \varepsilon$ となる事象であり，$|x_1+\cdots+x_n| \geqslant n\varepsilon$ とも書ける．ゆえに，

$$|x_1+\cdots+x_n| > \left(\frac{\varepsilon \cdot 2^{\alpha-1}}{c_{2^\alpha}}\right) e_{2^\alpha}.$$

ただし $c_{2^\alpha}^2$ は **4.5.1** で $\sigma_1^2+\cdots+\sigma_{2^\alpha}^2$ として定義されている．

しかし，コルモゴロフの不等式により次のようになる．

$$P(F_\alpha) \leqslant \frac{4c_{2^\alpha}^2}{\varepsilon^2 2^{2\alpha}}.$$

よって

(4.6.7)　$$\sum_{\alpha=1}^{\infty} P(F_\alpha) \leqslant \frac{4}{\varepsilon^2} \sum_{\alpha=1}^{\infty} 2^{-2\alpha} \sum_{i=1}^{2^\alpha} \sigma_i^2.$$

i と α の加算の順序を逆にし，$2^\alpha \geqslant i$ となるすべての正整数 α に関する和 $\sum 2^{-2\alpha}$ は $2i^{-2}$ を超えないから，

$$\sum_{\alpha=1}^{\infty} P(F_\alpha) \leqslant \frac{8}{\varepsilon^2} \sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2}$$

となる．ゆえに，$\sum_{i=1}^{\infty} \sigma_i^2/i^2$ が収束するならば，$\sum_{\alpha=1}^{\infty} P(F_\alpha)$ も収束する．この場合，十分に大きな α を選べばすべての値 β に対して，(4.6.6) の右辺を $<\delta$ とすることができる．十分に大きい N に対して，(4.6.4) と (4.6.1) が成り立つのは，(4.6.5) と (4.6.6) からわかる．これで **4.6.1** の証明が終わった．

$(x_1, x_2, \cdots)$ のすべての成分が等しい（有限な）分散を持てば，条件 $\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2}<+\infty$ は自動的に満たされることがいえる．実際，$x_1, x_2, \cdots$ が独立であり，平均0を持ち，同一に分布していれば，$(\bar{x}_1, \bar{x}_2, \cdots)$ は0に概収束する．Khintchine (1929) を参照せよ．

問　　題

4.1　$(x, y\ ;\ x_1, y_1\ ;\ x_2, y_2\ ;\ \cdots)$ が，$(x_1, x_2, \cdots)$，$(y_1, y_2, \cdots)$ がそれぞれ同値な確率変数 x, y に確率収束する確率過程であるとき，任意の $\varepsilon>0$ に対して次式が成り立つことを示せ．

$$\lim_{n\to\infty} P(|x_n-y_n|<\varepsilon)=1.$$

4.2（続き）$\lim_{n\to\infty} \mathscr{E}(x_n-y_n)^2=0$ かつ $(x_1, x_2, \cdots)$ が確率変数 x に確率収束するとき，$(y_1, y_2, \cdots)$ もまた x に確率収束することを示せ．

4.3　$(x_1, x_2, \cdots)$ が同一の c.d.f $F(x)$ を持つ独立な確率変数であり

$$F(x)=\begin{cases} 0, & x \leqslant 0 \\ x, & 0<x \leqslant 1 \\ 1, & x>1 \end{cases}$$

とする．また $(u_1, u_2, \cdots)$ を

$$u_n=\max(x_1, \cdots, x_n)$$

なる確率過程とし，$v_n=n(1-u_n)$ とすると，このとき列 $(v_1, v_2, \cdots)$ は $F(v)$ に法則収束することを示せ．

ただし

$$F(v)=\begin{cases} 1-e^{-v} & (v>0 \text{ に対して}) \\ 0 & (v \leqslant 0 \text{ に対して}). \end{cases}$$

4.4（続き）$g(x)$ は区間 $(0, \infty)$ 上の x について微分可能で連続な1価の増加関数であり，かつ $w_n=g(n(1-u_n))$ であるとき，$(w_1, w_2, \cdots)$ は $H(w)$ に法則収束することを示せ．ただし区間 $(g(0), g(+\infty))$ の w に対して，$\frac{dH(w)}{dw}$ は $e^{-g^{-1}(w)}\frac{d}{dw}g^{-1}(w)$ で与えられ，それ以外では0である．また $g^{-1}(w)$ は $g(x)$ の逆関数を示す．

4.5　離散型確率変数 x_n は次の c.d.f. を持つことがわかっているとする.

$$F_n(x) = 1 - \frac{n(n-1)\cdots(n-r+1)}{(n+nx)(n+nx-1)\cdots(n+nx-r+1)}.$$

ただし $r \leqslant n+1$, x_n の質点は $1/n, 2/n, \cdots$ である. このとき確率変数列 $(x_1, x_2, \cdots)$ が $F(x)$ に法則収束することを示せ. ただし

$$F(x) = \begin{cases} 0, & x \leqslant 0 \\ 1 - \dfrac{1}{(1+x)^r}, & x > 0. \end{cases}$$

4.6　**4.3.6 a** に関して, d が任意の定数 $(\fallingdotseq 0)$ のとき

$$\lim_{n \to \infty} P(x_n + ny_n \leqslant d) = 0$$

となることを示せ.

4.7　$(x_1, x_2, \cdots)$ は $\mathscr{E}(x_n) = 0$, $\sigma^2(x_n) = \sigma^2$, $n = 1, 2, \cdots$ かつ $\operatorname{cov}(x_i, x_j) = \rho\sigma^2$, $i \fallingdotseq j = 1, 2, \cdots$ となる確率過程とする. ただし σ^2 は有限で, ρ は $(0, 1)$ の間の値である. また $k = 1, 2, \cdots$ に対して

$$u_k = \frac{1}{k}(x_1 + x_3 + \cdots + x_{2k-1}), \qquad v_k = \frac{1}{k}(x_2 + x_4 + \cdots + x_{2k}).$$

とする. このとき, ベクトル値確率過程 $(u_1, v_1 ; u_2, v_2 ; \cdots)$ が (u, v) に法則収束することを示せ. しかも u, v は平均 0, 分散 $\sigma^2\rho$ を持つ同値な確率変数である.

4.8　$(x_1, \cdots, x_n)$ は n 個の独立変数で, すべて平均 0, 分散 1 を持つ. 次の不等式

$$|\bar{x}_\alpha| < \lambda\frac{1}{\sqrt{\alpha}}, \qquad \alpha = 1, \cdots, n$$

について考えよう. ただし

$$\bar{x}_\alpha = \frac{1}{\alpha}(x_1 + \cdots + x_\alpha).$$

G_n はすべての不等式を満たす事象とする. このとき $\overline{G_n}$ は少なくとも 1 つの不等式を満足しない事象である. $\overline{G_n}$ が起ったという条件下で, α 番目の不等式が最初に満たされないとするとき, コルモゴロフの不等式と同じ方法で, 任意の $\lambda > 0$ に対して次式が成り立つことを示せ.

$$P(\overline{G_n}) \leqslant \frac{n}{n + (\lambda^2 - 1)\mathscr{E}(\alpha|\overline{G_n})}.$$

ただし $\mathscr{E}(\alpha|\overline{G_n})$ は $\overline{G_n}$ が与えられたときの α の条件つき平均値であり, α は上の不等式を満たさない最初のものを示す確率変数を表わす.

4.9（続き）　次の式を考えよう.

$$x_1^2 + \cdots + x_\alpha^2 < \alpha\lambda^2, \ \alpha = 1, \cdots, n.$$

$\overline{G_n}$ が少なくとも 1 つの不等式を満たさない事象のとき, $P(\overline{G_n})$ は前記の問題と同じような不等式を満足することを示せ.

4.10　$(x_1, \cdots, x_n)$ が独立な正の確率変数で, すべての変数が平均 μ を持つとする. このとき, 任意の $\lambda > 0$ に対して, 不等式 $(x_1 x_2 \cdots x_\alpha) < \lambda\mu^\alpha$, $\alpha = 1, \cdots, n$ がすべて成り立つ確率は少なくとも $1 - 1/\lambda$ であることを示せ.

4.11　(有限な) 平均 $\mu_1, \mu_2, \cdots$ と (有限な) 分散 $\sigma_1^2, \sigma_2^2, \cdots$ を持つ独立な確率変数列 $(x_1, x_2, \cdots)$ に対して大数の強法則を拡張せよ.

4.12　$(x_1, \cdots, x_n)$ がすべてメディアン 0 を持つ n 個の独立な確率変数のとき, 次式を示せ.

$$\mathscr{E}(|x_1 + \cdots + x_n|) \geqslant \varphi(n)\overline{d}.$$

ただし

$$\overline{d} = \frac{1}{n}[\mathscr{E}|x_1| + \cdots + \mathscr{E}|x_n|]$$

かつ

$$\varphi(2k+1) = \varphi(2k+2) = \frac{(2k+1)!}{2^{2k}(k!)^2}.$$

これは Tukey (1946) による.

4.13　$(x_1, \cdots, x_k)$ は平均 $(0, \cdots, 0)$ と共分散行列 $\|\sigma_{ij}\|$ (ここで $\sigma_{ii} = \sigma^2$, $\sigma_{ij} = \sigma^2\rho$) を持つ k 次元確率変数とする. 次式を示せ.

$$P(\max_\alpha |x_\alpha| \geqslant t\sigma) \leqslant \frac{\sigma^2}{kt^2}[(k-1)\sqrt{1-\rho} + \sqrt{1+(k-1)\rho}].$$

これは Olkin と Pratt (1958) による.

4.14　2 つの確率過程 $(x_1, x_2, \cdots)$, $(y_1, y_2, \cdots)$ がそれぞれ確率変数 x, 正の定数 c に収束するならば, **4.3.6 a** で示された 4 つの式 (4.3.18), (4.3.19), (4.3.20), (4.3.21) が正しいことを示せ.

4.15　$(x_1, x_2, \cdots)$ が p.d.f. $(f_1(x_1), f_2(x_2), \cdots)$ を持つ確率変数列で, かつ確率 0 の集合をのぞいた R_1 のすべての x に対して $\lim_{\alpha \to \infty} f_\alpha(x) = f(x)$ であれば,

$$\lim_{\alpha \to \infty} \int_E f_\alpha(x)\,dx = \int_E f(x)\,dx$$

は $E \subset R_1$ のすべての集合で一様である. ただし $f(x)$ は p.d.f. である. [この結果は Scheffé (1947) による. シェフィーの定理におけるこの条件は **4.3.4** の条件よりも強い. 法則収束のための条件の議論および比較は Robbins (1948) に述べられている].

4.16　$(x, x_1, x_2, \cdots)$ は $(x_1, x_2, \cdots)$ が確率変数 x に確率収束する確率過程であり, $g_1(x), g_2(x), \cdots$ は有界区間上で $g(x)$ に一様収束する連続な関数列である. このとき $(g_1(x_1), g_2(x_2), \cdots)$ は $g(x)$ に確率収束することを示せ

4.17　$x_1, \cdots, x_n$ を独立な確率変数とし, $y_1 = x_1, y_2 = x_1 + x_2, \cdots, y_n = x_1 + \cdots + x_n$ とする. y_ξ の c.d.f. を $F_\xi(y_\xi)$, $\xi = 1, \cdots, n$ で表わし, $(y_1, \cdots, y_n)$ の c.d.f. を $F(y_1, \cdots, y_n)$ で表わすと次式が成立することを示せ.

$$F(y_1, \cdots, y_n) \geqslant F_1(y_1) \cdots F_n(y_n).$$

これは Robbins (1954) による.

4.18　**4.3.7** を証明せよ.

第5章　特性関数と母関数

5.1　1 変数の場合

数理統計学で重要な問題の 1 つに確率変数の可測関数の分布関数，すなわち確率変数の関数で，かつそれ自身が確率変数であるような関数の分布関数を決定する問題がある．この種の問題の取り扱いに関してのいくつかの手法を 2.8 節で述べた．しかし，この手法を個々の場合に用いるとき，技術的な難点や繁雑さが生じて来る．ある状態，特に独立な確率変数の線形関数を含む場合には，対象として考察しようとする確率変数の特性関数を使用すると，エレガントに取り扱いうることがある．特性関数とそれに関連した手法は分布の母積率や累積率[*]の統計的処理，または確率変数の 2 つあるいはそれ以上の関数の独立性を検定する際には有用であろう．この章ではこの方法とその応用について述べよう．

c. d. f. $F(x)$ を持つ確率変数の**特性関数** $\varphi(t)$ は次式で定義される．

(5. 1. 1)
$$\varphi(t)=\mathscr{E}(e^{\mathrm{i}tx})=\int_{-\infty}^{\infty}e^{\mathrm{i}tx}\,dF(x).$$

ここで $\mathrm{i}=\sqrt{-1}$ で，t は実数である[**]．便宜上，$\varphi(t)$ を $F(x)$ に対応する特性関数，あるいは x の特性関数と呼ぶ．

$$e^{\mathrm{i}tx}=\cos tx+\mathrm{i}\sin tx$$

だから，$\cos tx$ と $\sin tx$ はともに任意の t に対して，R_1 上で積分可能であるので，$\varphi(t)$ は複素数であり，その実数部分と虚数部分は t の任意の値に対して有限である．

モーメント母関数 $\psi(t)$ は次の式で定義される．

(5. 1. 2)
$$\psi(t)=\mathscr{E}(e^{tx})=\varphi\left(\frac{t}{\mathrm{i}}\right).$$

---

[*] 半不変係数．（訳注）

[**] この本では，他章で i を総和の添字として使用したので，混乱をさけるため i で $\sqrt{-1}$ を表わす．

$t=0$ のある近傍における t の値に対する $\psi(t)$ が通常問題になる．しかし $\psi(t)$ はすべての c. d. f. に対して存在するとは限らないが，t のすべての値に対して $\varphi(t)$ は存在する．

階乗モーメント母関数 $\theta(t)$ はもしそれが存在すれば，(5. 1. 2) 式において t を $\log t$ に置き換えることにより得られる．つまり

(5. 1. 3)
$$\theta(t)=\psi(\log t)=\mathscr{E}(t^x).$$

x がその質点が正の整数であるような離散型確率変数のとき，$\theta(t)$ が存在すれば，確率変数 x の**確率母関数**とも呼ばれる．なぜなら $\theta(t)$ が t に関して級数に展開できれば，t^x の係数は x の p. f. $p(x)$ であるからである．

r 次のモーメント $\mu'_r(x)$ が存在すれば，$0<h\leqslant r$ なる h に対して，t に関して (5. 1. 1) 式を h 回微分でき，次式を得る．

(5. 1. 4)
$$\varphi^{(h)}(t)=\mathrm{i}^h\int_{-\infty}^{\infty}x^h e^{\mathrm{i}tx}\,dF(x).$$

したがって h 次モーメント $\mu'_h(x)$ は次のようになる．

(5. 1. 5)
$$\mu'_h(x)=\frac{\varphi^{(h)}(0)}{\mathrm{i}^k},\qquad 0<h\leqslant r.$$

$\varphi^{(0)}(t)$ を $\varphi(t)$ と簡単に記す．

同様に，0 の近傍における t の値に対して $\psi(t)$ が存在すれば次式を得る．

(5. 1. 6)
$$\mu'_h(x)=\psi^{(h)}(0).$$

x の $2s$ 次のモーメントまでのすべてのモーメントが存在すれば次のように書ける．

(5. 1. 7)
$$\varphi(t)=\int_{-\infty}^{\infty}\left[1+\sum_{h=1}^{2s-1}\frac{(\mathrm{i}tx)^h}{h!}+\frac{(\mathrm{i}tx)^{2s}}{(2s)!}(\cos t'x+\mathrm{i}\sin t''x)\right]dF(x).$$

ただし t' と t'' は区間 $(0,t)$ 内の数である．x が非退化であるときは $\mu'_{2s}\neq 0$ であり，次式のように書ける．

(5. 1. 8)
$$\varphi(t)=1+\sum_{h=1}^{2s-1}\frac{(\mathrm{i}t)^h}{h!}\mu'_h+\frac{(\mathrm{i}t)^{2s}\mu'_{2s}}{(2s)!}g_1(t,s).$$

ここで $g_1(t,s)$ は複素関数であり，実数部分と虚数部分はそれぞれ $\mathscr{E}(x^{2s}\cos t'x/\mu'_{2s})$ と $\mathscr{E}(x^{2s}\sin t''x/\mu'_{2s})$ なる平均値である．ただし $|g_1(t,s)|\leqslant 1$ である．さらに，t' と t'' の両方とも区間 $(0,t)$ 内に存在するので，$\lim_{t\to 0}g_1(t,s)=1$ である．

同様に，もし $(2s-1)$ 次のモーメントを含めて $(2s-1)$ 次以下のすべてのモーメントが存在すれば

(5. 1. 9)
$$\varphi(t)=1+\sum_{h=1}^{2s-2}\frac{(\mathrm{i}t)^h}{h!}\mu'_h+\frac{(\mathrm{i}t)^{2s-1}}{(2s-1)!}\nu'_{2s-1}g_2(t,s)$$

と書ける．ここで $g_2(t, s)$ は，実数部分が $\mathscr{E}(x^{2s-1}\cos t_1 x/\nu'_{2s-1})$ であり，虚数部分が $\mathscr{E}(x^{2s-1}\sin t_2 x/\nu'_{2s-1})$ となる複素関数であり，t_1 と t_2 は区間 $(0, t)$ の間の数である．また ν'_{2s-1} は，x が非退化であれば (3.3.13) で定義された $\neq 0$ の $(2s-1)$ 次の絶対モーメントである．ただし $|g_2(t, s)| \leqslant 1$ となる．さらに $\lim_{t\to 0} g_2(t, s) = \dfrac{\mu'_{2s-1}}{\nu'_{2s-1}}$ はもちろん有限である．

最高次のモーメントが偶数次，奇数次のいずれでも，まとめると次のようになる．

5.1.1　確率変数 x の r 次のモーメントが存在すれば，$\varphi(t)$ は $t=0$ の近傍で次のように展開できる．

(5.1.10)
$$\varphi(t) = 1 + \sum_{h=1}^{r} \frac{(it)^h}{h!}\mu'_h + o(t^r)$$

ただし
$$\lim_{t\to 0}\frac{o(t^r)}{t^r} = 0.$$

$\varphi(t)$ が (5.1.10) に展開可能ならば，$\log\varphi(t)$ もまた次のように展開される．

(5.1.11)
$$\log\varphi(t) = \sum_{h=1}^{r}\frac{\kappa_h}{h!}(it)^h + o(t^r).$$

値 κ_h は c.d.f. $F(x)$ の半不変係数または累積率と呼ばれ，Thiele (1903) が最初に定義した．

どのような κ_h もモーメント $\mu'_1, \mu'_2, \cdots$ における多項式である．最初の 2，3 の半不変係数は次のようになる．

(5.1.12)
$$\begin{aligned}\kappa_1 &= \mu'_1 = \mu\\ \kappa_2 &= \mu'_2 - (\mu'_1)^2 = \sigma^2\\ \kappa_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3\\ &\vdots\end{aligned}$$

逆に

(5.1.13)
$$\begin{aligned}\mu'_1 &= \kappa_1\\ \mu'_2 &= \kappa_2 + \kappa_1^2\\ \mu'_3 &= \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3\\ &\vdots\end{aligned}$$

確率論とその応用，特に標本論には多くの問題がある．標本論では多次元確率変数の成分の関数，特に線形関数の特性関数は容易に求められる．もちろんここで特性関数から確

率変数の c.d.f. をいかにして見つけるかという問題が起る．これは次に述べる Lévy (1925) の定理から，多くの場合に可能となる．

5.1.2　x を特性関数 $\varphi(t)$ と c.d.f. $F(x)$ を持つ確率変数とする．このとき $F(x)$ が $x = x' \pm \delta$, $\delta > 0$ で連続であれば

(5.1.14)
$$F(x'+\delta) - F(x'-\delta) = \lim_{A\to\infty}\frac{1}{\pi}\int_{-A}^{A}\frac{\sin\delta t}{t}e^{-itx'}\varphi(t)\,dt$$

である．さらに，$\int_{-\infty}^{\infty}|\varphi(t)|\,dt < +\infty$ であれば，$x = x'$ で p.d.f. $f(x)$ が存在し

(5.1.15)
$$f(x') = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx'}\varphi(t)\,dt$$

となる．

この定理の証明は，まず (5.1.14) における $\varphi(t)$ を (5.1.1) における積分表現で置き換える．まず

(5.1.16)
$$G(x', A, \delta) = \frac{1}{\pi}\int_{-A}^{A}\frac{\sin\delta t}{t}e^{-itx'}\int_{-\infty}^{\infty}e^{itx}dF(x)\,dt$$

と置こう．ここで $\left|\dfrac{\sin\delta t}{t}e^{-it(x'-x)}\right| < \delta$ であるので，(5.1.16) の積分の順序を変えることができるから

(5.1.17)
$$G(x', A, \delta) = \int_{-\infty}^{\infty} m(x, x', A, \delta)\,dF(x)$$

となる．さらに

(5.1.18)
$$\begin{aligned}m(x, x', A, \delta) &= \frac{1}{\pi}\int_{-A}^{A}\frac{\sin\delta t}{t}e^{it(x-x')}\,dt\\ &= \frac{2}{\pi}\int_{0}^{A}\frac{\sin\delta t}{t}\cos[t(x-x')]\,dt\\ &= \frac{1}{\pi}\int_{0}^{A}\frac{\sin(x-x'+\delta)t}{t}\,dt\\ &\quad - \frac{1}{\pi}\int_{0}^{A}\frac{\sin(x-x'-\delta)t}{t}\,dt\end{aligned}$$

である．極限をとると

(5.1.19)
$$\lim_{A\to\infty}G(x', A, \delta) = \int_{-\infty}^{\infty}\lim_{A\to\infty}m(x, x', A, \delta)\,dF(x).$$

しかし $\lim_{V\to\infty}\int_0^V \dfrac{\sin u}{u}\,du = \dfrac{\pi}{2}$ を使用すると，次のようになる．

$$\lim_{A\to\infty} m(x,x',A,\delta)=\begin{cases}0, & x<x'-\delta,\quad x>x'+\delta\\ 1/2, & x=x'-\delta,\quad x=x'+\delta\\ 1, & x'-\delta<x<x'+\delta.\end{cases} \tag{5.1.20}$$

したがって，$F(x)$ は $x=x'\pm\delta$ で連続なので，(5.1.19) の $\lim_{A\to\infty} m(x,x',A,\delta)$ の値を使用すると

$$\lim_{A\to\infty} G(x',A,\delta)=\int_{x'-\delta}^{x'+\delta} dF(x)=F(x'+\delta)-F(x'-\delta)$$

となり，(5.1.14) が証明できた.

次に，(5.1.14) を 2δ で割れば

$$\frac{F(x'+\delta)-F(x'-\delta)}{2\delta}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\sin\delta t}{\delta t}e^{-itx'}\varphi(t)\,dt \tag{5.1.21}$$

となる. $F(x)$ が $x=x'$ で導関数 $f(x')$ を持てば，$\left|\frac{\sin\delta t}{\delta t}e^{-itx'}\varphi(t)\right|<|\varphi(t)|$ だから，$\int_{-\infty}^{\infty}|\varphi(t)|dt<+\infty$ であれば

$$\lim_{\delta\to 0}\left[\frac{F(x'+\delta)-F(x'-\delta)}{2\delta}\right]=\frac{1}{2\pi}\int_{-\infty}^{\infty}\lim_{\delta\to 0}\left(\frac{\sin\delta t}{\delta t}\right)e^{-itx'}\varphi(t)\,dt$$

となる. ここで $\lim_{\delta\to 0}(\sin\delta t/\delta t)=1$ なので

$$f(x')=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx'}\varphi(t)\,dt$$

となり，これは (5.1.15) となる.

2つの確率変数 x_1 と x_2 が同じ c.d.f. を持てば，明らかにそれらの特性関数は等しい. 逆に確率変数が同じ特性関数 $\varphi(t)$ を持ったとする. このとき $F_1(x)$，$F_2(x)$ が2つの確率変数の c.d.f. ならば，(5.1.14) から次のことが示される. すなわち，$x'\pm\delta$ が $F_1(x)$ と $F_2(x)$ が端点 $x'\pm\delta$ で連続となる任意の区間ならば

$$F_1(x'+\delta)-F_1(x'-\delta)=F_2(x'+\delta)-F_2(x'-\delta)$$

となり，したがって $F_1(x)$ と $F_2(x)$ は c.d.f. であり，すべての c.d.f. は，右連続であるということを考慮すれば，$F_1(x)\equiv F_2(x)$ となる.

これにより次の結果を得る.

5.1.3　x_1，x_2 がそれぞれ c.d.f. $F_1(x)$，$F_2(x)$ および特性関数 $\varphi_1(t)$，$\varphi_2(t)$ を持つ確率変数のとき，$F_1(x)\equiv F_2(x)$ となるための必要十分条件は $\varphi_1(t)\equiv\varphi_2(t)$ である.

この c.d.f. と特性関数との間の1対1対応は c.d.f. が対応する特性関数と同一視されるから，確率論では非常に有用である.

例題　$x_1,\cdots,x_k$ は統計的に独立な確率変数であり，p.d.f.

$$f(x_1,\cdots,x_k)=\begin{cases}e^{-\sum_1^k x_i}, & x_i\geqq 0,\ i=1,\cdots,k\\ 0, & \text{その他}\end{cases}$$

を持つ. $L=\sum_1^k x_i$ のとき，L のモーメントと p.d.f. を求めよう.

L の特性関数は次式となる.

$$\varphi(t)=\int_0^{\infty}\cdots\int_0^{\infty}e^{-\sum_i x_i+it\sum_i x_i}\,dx_1\cdots dx_k=(1-it)^{-k}. \tag{5.1.22}$$

r 次のモーメント $\mu_r'(L)$ は (5.1.5) を適用すると

$$\mu_r'(L)=\frac{\varphi^{(r)}(0)}{i^r}=\frac{\Gamma(k+r)}{\Gamma(k)}$$

として与えられる. L の p.d.f. は (5.1.15) により

$$f(L)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itL}(1-it)^{-k}\,dt \tag{5.1.23}$$

となる. 積分は $z=-L(1-it)$ なる変数変換をして，複素平面における等高積分*) によって評価できる.

(5.1.23) の積分は $L^{k-1}e^{-L}\cdot H$ となる. ただし H はハンケル積分で

$$H=\frac{i}{2\pi}\int_{-L+i\infty}^{-L-i\infty}e^{-z}(-z)^{-k}\,dz \tag{5.1.24}$$

で表わされる. この値は $1/\Gamma(k)$ であることがわかる. [たとえば Whittaker と Watson (1927) を参照せよ]. L の p.d.f. はそれゆえ

$$f(L)=\begin{cases}\frac{1}{\Gamma(k)}L^{k-1}e^{-L}, & L\geqq 0\\ 0, & L<0\end{cases} \tag{5.1.25}$$

で与えられる.

5.2　k 変数の場合

$(x_1,\cdots,x_k)$ を c.d.f. $F(x_1,\cdots,x_k)$ を持つ k 次元確率変数とする. $(x_1,\cdots,x_k)$ の特性関数 $\varphi(t_1,\cdots,t_k)$ は

*) 原書では contour integration となっているが，ここでは等高積分とした.（訳注）

$$(5.2.1)\qquad \varphi(t_1,\cdots,t_k)=\mathscr{E}\left[\exp\left(\mathrm{i}\sum_i^k t_ix_i\right)\right]$$

$$=\int_{R_k}\exp\left(\mathrm{i}\sum_1^k t_ix_i\right)dF(x_1,\cdots,x_k)$$

で定義される.

モーメント母関数 $\psi(t_1,\cdots,t_k)$ は

$$(5.2.2)\qquad \psi(t_1,\cdots,t_k)=\varphi\left(\frac{t_1}{\mathrm{i}},\cdots,\frac{t_k}{\mathrm{i}}\right)$$

で定義される.

同様に, 階乗モーメント母関数 $\theta(t_1,\cdots,t_k)$ は $\psi(t_1,\cdots,t_k)$ において, t_i を $\log t_i$ $(i=1,\cdots,k)$ で置き換えて得られる.

結合モーメント $\mu'_{r_1\cdots r_k}$ が存在すれば, $\varphi(t_1,\cdots,t_k)$ を微分して次のように得られる. すなわち

$$(5.2.3)\qquad \mu'_{r_1\cdots r_k}=\frac{\varphi^{(r_1+\cdots+r_k)}(0,\cdots,0)}{\mathrm{i}^{(r_1+\cdots+r_k)}}.$$

$\mu'_{r_1\cdots r_k}$ が存在すれば, すべての結合モーメント $\mu'_{h_1\cdots h_k}$, $0<h_i\leqslant r_i$, $i=1,\cdots,k$ が存在する.

確率変数 $(x_1,\cdots,x_k)$ の成分の任意の部分集合の特性関数は, 部分集合に含まれない確率変数に対応する t の値を 0 と置くことにより得られる. たとえば確率変数 $(x_1,\cdots,x_{k_1})$, $k_1<k$ の特性関数は $\varphi(t_1,\cdots,t_{k_1},0,\cdots,0)$ である.

レヴィの定理の k 次元確率変数への拡張は次のようになる.

5.2.1　$(x_1,\cdots,x_k)$ を特性関数 $\varphi(t_1,\cdots,t_k)$ と c.d.f. $F(x_1,\cdots,x_k)$ を持つ k 次元確率変数とする. I_k を R_k における区間 $x'_i-\delta_i<x_i\leqslant x'_i+\delta_i$, $i=1,\cdots,k$, $\delta_i>0$, とする. $F(x_1,\cdots,x_k)$ を閉区間 $x'_i-\delta_i\leqslant x_i\leqslant x'_i+\delta_i$, $i=1,\cdots,k$ の境界上で連続とする. このとき

$$(5.2.4)\qquad P((x_1,\cdots,x_k)\in I_k)$$

$$=\lim_{A\to\infty}\frac{1}{\pi^k}\int_{-A}^{A}\cdots\int_{-A}^{A}\prod_{i=1}^{k}\left[\frac{\sin\delta_it_i}{t_i}e^{-\mathrm{i}t_ix'_i}\right]\varphi(t_1,\cdots,t_k)\,dt_1\cdots dt_k$$

となり, さらに $\int_{R_k}|\varphi(t_1,\cdots,t_k)|dt_1\cdots dt_k<+\infty$ ならば $(x'_1,\cdots,x'_k)$ で p.d.f. $f(x_1,\cdots,x_k)$ が存在し,

$$(5.2.5)\qquad f(x'_1,\cdots,x'_k)$$

$$=\left(\frac{1}{2\pi}\right)^k\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left(-\mathrm{i}\sum_1^k t_ix'_i\right)\varphi(t_1,\cdots,t_k)\,dt_1\cdots dt_k$$

が成り立つ.

5.2.1 の証明は **5.1.2** と同様であり, 省略する.

c.d.f. と特性関数との間の 1 対 1 対応に関する **5.1.3** は k 次元確率変数の場合へも容易に拡張できる.

5.3　独立変数の特性関数

特性関数は, 確率変数の分布関数を求めなくても, 単に確率変数が独立かどうかをこれによって決定することが出来るので有用である. これを次に述べる.

5.3.1　(x_1,x_2) が 2 次元確率変数ならば, x_1 と x_2 が独立であるための必要十分条件は

$$(5.3.1)\qquad \varphi(t_1,t_2)=\varphi(t_1,0)\cdot\varphi(0,t_2)$$

である. ただし $\varphi(t_1,t_2)$ は (x_1,x_2) の特性関数である.

ここで $\varphi(t_1,0)$ は x_1 の, $\varphi(0,t_2)$ は x_2 の特性関数であり, 簡単に $\varphi_1(t_1)$, $\varphi_2(t_2)$ と書く.

必要性を見るために, x_1 と x_2 は独立であると仮定する. すなわち $F(x_1,x_2)=F_1(x_1)F_2(x_2)$ とする. よって

$$(5.3.2)\qquad \varphi(t_1,t_2)=\int_{R_2}e^{\mathrm{i}t_1x_1+\mathrm{i}t_2x_2}dF(x_1,x_2)$$

$$=\int_{R_2}e^{\mathrm{i}t_1x_1+\mathrm{i}t_2x_2}d[F_1(x_1)F_2(x_2)]$$

$$=\int_{-\infty}^{\infty}e^{\mathrm{i}t_1x_1}dF_1(x_1)\int_{-\infty}^{\infty}e^{\mathrm{i}t_2x_2}dF_2(x_2)$$

となる. すなわち

$$\varphi(t_1,t_2)=\varphi_1(t_1)\cdot\varphi_2(t_2).$$

十分性について考える. (5.3.1) が成り立つと仮定して, $k=2$ に対して (5.2.4) を

適用すると

(5.3.3)　$P(x_i' - \delta_i < x_i \leqslant x_i' + \delta_i\,;\, i = 1, 2)$

$$= \lim_{A\to\infty} \frac{1}{\pi^2} \int_{-A}^{A} \int_{-A}^{A} \prod_{i=1}^{2} \left[\frac{\sin \delta_i t_i}{t_i} e^{-\mathrm{i} t_i x_i'} \right] \varphi_1(t_1)\varphi_2(t_2)\, dt_1\, dt_2$$

$$= \prod_{i=1}^{2} \left[\lim_{A\to\infty} \frac{1}{\pi} \int_{-A}^{A} \frac{\sin \delta_i t_i}{t_i} e^{-\mathrm{i} t_i x_i'} \varphi_i(t_i)\, dt_i \right]$$

となる．すなわち

(5.3.4)　$$P((x_1, x_2) \in I_2) = P(x_1 \in I_1^{(1)}) \cdot P(x_2 \in I_1^{(2)}).$$

ただし $I_1^{(i)}$ は区間 $x_i' - \delta_i < x_i \leqslant x_i' + \delta_i$，$i = 1, 2$ であり，I_2 は直積 $I_1^{(1)} \times I_2^{(2)}$ である．しかし **2.4.2** より (5.3.4) は式

$$F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$$

である．すなわち x_1 と x_2 は独立である．

5.3.1 が任意の（有限）個の確率変数へ拡張可能なことは明白であり，読者に残す．

独立な確率変数の特性関数については，確率変数が線形関数の場合には，特に有用な結果が得られる．この問題に関する結果を次に示すが，読者は容易に確かめることができよう．

5.3.2　L を次式で定義された k 個の独立な確率変数 $x_1, \cdots, x_k$ の線形関数とする．

(5.3.5)　$$L = \sum_{i=1}^{k} c_i x_i.$$

$\varphi(t)$ を L の特性関数とし，$\varphi_i(t_i)$ を x_i $(i = 1, \cdots, k)$ の特性関数とする．このとき

(5.3.6)　$$\varphi(t) = \prod_{i=1}^{k} \varphi_i(c_i t)$$

が成り立つ．

x_1，x_2 は独立な確率変数であり，c.d.f. $F(x_1\,;\,\theta_1)$，$F(x_2\,;\,\theta_2)$ を持つとする．ただし θ_1 と θ_2 はパラメータ θ の値である．確率変数 $x_1 + x_2$ を L と書き，L の c.d.f. が $F(L\,;\,\theta_1 + \theta_2)$ であれば，c.d.f. $F(x\,;\,\theta)$ は θ に関して**再生的**であるという．同様に θ に依存する p.f.，p.d.f.，分布についても θ に関して再生的であるということができる．再生性は重要な性質で，後章でしばしば用いられる．再生性の概念は確率変数 x とパラメータ θ のどちらか一方または両方が多次元になった場合にも容易に拡張できる．特性関数は次に述べるように，c.d.f. $F(x\,;\,\theta)$ が再生的であるかどうかを決定するための有用な基準になっている．

5.3.3　x_1，x_2 は独立な確率変数であり，おのおのが c.d.f. $F(x_1\,;\,\theta_1)$，$F(x_2\,;\,\theta_2)$，および特性関数 $\varphi(t\,;\,\theta_1)$，$\varphi(t\,;\,\theta_2)$ を持つとする．ただし θ_1，θ_2 はパラメータである．このとき $F(x\,;\,\theta)$ が θ に関して再生的であるための必要十分条件は

(5.3.7)　$$\varphi(t\,;\,\theta_1)\varphi(t\,;\,\theta_2) = \varphi(t\,;\,\theta_1 + \theta_2)$$

である．

この証明は簡単であるので読者にまかそう．

5.4　確率変数列の特性関数

4.1 節で述べたように，確率論およびその標本論への応用での重要な問題の 1 つとして，確率変数列の確率収束に関するものがある．c.d.f. は特性関数により一意に定まるので，この問題は確率変数の c.d.f. を直接扱うよりも，対応する特性関数列の収束問題として扱った方が容易である．この基本的な原理は Lévy (1937) と Cramér (1937) による次の定理で示される．

5.4.1　$(x_1, x_2, \cdots)$ を確率変数列とし，$\varphi_1(t), \varphi_2(t), \cdots$ を対応する特性関数列とする．$(x_1, x_2, \cdots)$ が c.d.f. $F(x)$ に法則収束するための必要十分条件は t のどの値に対しても，列 $\varphi_1(t), \varphi_2(t), \cdots$ が $t = 0$ で連続な極限 $\varphi(t)$ に収束することである．これらの条件の下で，$\varphi(t)$ は対応する特性関数 $F(x)$ と同一視できる．

まず必要性については，$F_1(x), F_2(x), \cdots$ が $(x_1, x_2, \cdots)$ の c.d.f. であり，$F(x)$ が c.d.f. の極限であるとき，列 $F_1(x), F_2(x), \cdots$ は各 x に対して c.d.f. $F(x)$ に収束すると仮定する．このとき任意の t に対して $\lim_{n\to\infty} \varphi_n(t) = \varphi(t)$ となることを示す．

(5.4.1)　$$\varphi_n(t) = \int_{-\infty}^{\infty} \cos tx\, dF_n(x) + \mathrm{i} \int_{-\infty}^{\infty} \sin tx\, dF_n(x)$$

のように書ける．$\cos tx$ は任意の t に対して $(-\infty, +\infty)$ 上で有界であり，$F_1(x)$, $F_2(x), \cdots$ は任意の x に対して c.d.f. $F(x)$ に収束するので，ヘリー=ブレイの定理により［たとえば Loève (1955) を参照せよ］，

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\cos tx\,dF_n(x)=\int_{-\infty}^{\infty}\cos tx\,dF(x) \tag{5.4.2}$$

が成り立ち，同様に

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\sin tx\,dF_n(x)=\int_{-\infty}^{\infty}\sin tx\,dF(x) \tag{5.4.3}$$

である．

しかし，(5.4.2) と (5.4.3) をまとめると

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}e^{itx}\,dF_n(x)=\int_{-\infty}^{\infty}e^{itx}\,dF(x)$$

と同値である．すなわち

$$\lim_{n\to\infty}\varphi_n(t)=\varphi(t)$$

である．

十分性に関しては，任意の t に対して $\lim_{n\to\infty}\varphi_n(t)=\varphi(t)$ および $t=0$ で $\varphi(t)$ は連続であると仮定する．このとき c.d.f. の列 $F_1(x), F_2(x), \cdots$ の中から右連続，非減少関数 $F(x)$ に収束する部分列 $F_{n_1}(x), F_{n_2}(x), \cdots$ を選ぶことが示せる．［たとえば Cramér (1946) を参照せよ］．ここで問題は，部分列 $F_{n_1}(x), F_{n_2}(x), \cdots$ が非減少，右連続な $F(x)$ に収束し，$F(x)$ が c.d.f. になるための他の条件，つまり $F(-\infty)=0$ と $F(+\infty)=1$ を満足していることを示せば良い．$0 \leqslant F(x) \leqslant 1$ は明らかである．

ここで (5.1.14) の場合と同じ方法を使う．すなわち $c>0$ に対して

$$c\left[\frac{1}{c}\int_0^c F_{n_i}(y)\,dy-\frac{1}{c}\int_{-c}^0 F_{n_i}(y)\,dy\right]=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos ct}{t^2}\varphi_{n_i}(t)\,dt \tag{5.4.4}$$

が成り立つ．(5.4.4) の左辺の第 1 項において，もし y が，区間 $(0,c)$ で p.d.f. $1/c$ を持つ連続型確率変数とすると，$F_{n_i}(y)$ はパラメータ n_i の関数である確率変数となることがわかる．$F_{n_1}(x), F_{n_2}(x), \cdots$ は $F(x)$ に収束し，これらすべての関数は区間 $[0,1]$ の値をとるので

$$\lim_{i\to\infty}\frac{1}{c}\int_0^c F_{n_i}(y)\,dy=\frac{1}{c}\int_0^c F(y)\,dy.$$

同様に

$$\lim_{i\to\infty}\frac{1}{c}\int_{-c}^0 F_{n_i}(y)\,dy=\frac{1}{c}\int_{-c}^0 F(y)\,dy.$$

同様な考え方で，(5.4.4) の右辺に $i\longrightarrow\infty$ なる極限をとると

$$c\left[\frac{1}{c}\int_0^c F(y)\,dy-\frac{1}{c}\int_{-c}^0 F(y)\,dy\right]=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos ct}{t^2}\varphi(t)\,dt. \tag{5.4.5}$$

$t=u/c$ と置き，(5.4.5) の両辺を c で割ると

$$\frac{1}{c}\int_0^c F(y)\,dy-\frac{1}{c}\int_{-c}^0 F(y)\,dy=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos u}{u^2}\varphi\left(\frac{u}{c}\right)du \tag{5.4.6}$$

となる．

$c\longrightarrow\infty$ とすると $F(y)$ は非減少なので，(5.4.6) の左辺の極限は $F(+\infty)-F(-\infty)$ となる．$\varphi(t)$ は $t=0$ で連続だから，任意の u に対して，$\lim_{c\to\infty}\varphi(u/c)=\varphi(0)$ となる．しかし任意の t に対して $\varphi_1(t), \varphi_2(t), \cdots$ は $\varphi(t)$ に収束する．ゆえに $\lim_{n\to\infty}\varphi_n(0)=\varphi(0)$ となる．しかも任意の n に対しても $\varphi_n(0)=1$ なので，$\varphi(0)=1$ となり，最後に $c\longrightarrow\infty$ として (5.4.6) から

$$F(+\infty)-F(-\infty)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos u}{u^2}\,du=1 \tag{5.4.7}$$

を得る．$F(x)$ は非負で，1 を越えない．よって，$F(+\infty)=1$，$F(-\infty)=0$ となる．ここで $F_{n_1}(x), F_{n_2}(x), \cdots$ の極限 $F(x)$ は c.d.f. となり，$F(x)$ の特性関数は $\varphi_{n_1}(t)$, $\varphi_{n_2}(t), \cdots$ の極限 $\varphi(t)$ になる．さらに，$F_1(x), F_2(x), \cdots$ の，非減少関数に収束する他の部分列があれば，この極限を $F^*(x)$ とする．

このとき前と同様に，$F^*(x)$ は c.d.f. であり，その特性関数は $\varphi(t)$ と同じになることが示せる．$F(x)$ と $F^*(x)$ は同じ特性関数を持つ c.d.f. なので，**5.1.3** より $F(x) \equiv F^*(x)$ である．よって $F_1(x), F_2(x), \cdots$ の任意の収束部分列は c.d.f. $F(x)$ に収束する．もちろん，これは $F_1(x), F_2(x), \cdots$ が c.d.f. $F(x)$ に収束するのと同値である．これで **5.4.1** の証明は終わった．

5.4.1 に関する次の系は定数に確率収束することを示す問題には有用である．

5.4.1 a　特性関数 $\varphi_1(t), \varphi_2(t), \cdots$ を持つ確率変数列 $x_1, x_2, \cdots$ が定数 c に確率収束するための必要十分条件は $\lim_{n\to\infty}\varphi_n(t)=e^{ict}$ である．

5.4.1 と **5.4.1 a** の k 次元への拡張は容易なので読者にゆだねる．

5.5　モーメントによる分布関数の決定

(a)　モーメント列による c.d.f. の決定

5.1 節では，分布関数のモーメントは，もし存在すれば，特性関数を微分することにより得られ，(5.1.10) は特性関数とそのモーメントの存在関係を示している．しかし，確率変数の c.d.f. のモーメントは特性関数を微分する以外の方法で求めることができる場合もある．確率変数および確率変数の関数のモーメント列が 1 つまたは別の方法でも容易に定められることが後の章で，しばしばでてくるであろう．しかしこれに関する基本的な問題は，次の通りである．いかなる条件の下で，確率変数 x のモーメント列 $\mu_0', \mu_1', \mu_2', \cdots$ が c.d.f. を一意に決定できるかということである．これに対する有用な十分条件は次の通りである [Cramér (1946) を参照せよ]．

5.5.1　$F(x)$ をすべてが有限なモーメント μ_r' $(r = 0, 1, 2, \cdots)$ を持つ c.d.f. とする．級数 $\sum_{r=0}^{\infty} \frac{\mu_r'}{r!} c^r$ が，ある $c > 0$ に対して絶対収束すれば，$F(x)$ はこれらのモーメントを持つただ 1 つの c.d.f. である．

$F(x)$ に関する特性関数の定義 (5.1.1) および (5.1.10) を導いたのと同様な方法によって

(5.5.1)
$$\varphi(t+u) = \int_{-\infty}^{\infty}\left[1 + \sum_{r=1}^{n-1} \frac{(\mathrm{i}u)^r}{r!} x^r + \frac{(\mathrm{i}u)^n}{n!} x^n (\cos u'x + \mathrm{i} \sin u''x)\right] e^{\mathrm{i}tx}\, dF(x)$$

と書ける．ただし u' と u'' は区間 $(0, u)$ の（実）数である．n 次モーメントが存在するならば，(5.1.4) を適用すると

(5.5.2)
$$\varphi(t+u) = \sum_{r=0}^{n-1} \frac{u^r}{r!} \varphi^{(r)}(t) + \frac{\nu_n' u^n}{n!} q$$

が得られる．ただし q は $|q| \leqslant 1$ なる複素関数であり，ν_n' は x の n 次絶対モーメントである．n が偶数であれば $\nu_n' = \mu_n'$ であり，定理の仮定により，剰余 $|u| < c$ ならば $(\nu_n' u^n / n!) q \longrightarrow 0$ $(n \longrightarrow \infty)$ となる．ここで n が奇数の剰余項ならば任意の λ に対して

(5.5.3)
$$\mathscr{E}[\lambda |x|^{(n-1)/2} + |x|^{(n+1)/2}]^2 = \lambda^2 \nu_{n-1}' + 2\lambda \nu_n' + \nu_{n+1}' \geqslant 0$$

である．これから

(5.5.4)
$$\nu_n'^2 \leqslant \nu_{n-1}' \cdot \nu_{n+1}'.$$

これは

(5.5.5)
$$\frac{\nu_n'}{n!} u^n \leqslant \left[\left(\frac{\nu_{n-1}'}{(n-1)!} u^{n-1}\right)\left(\frac{\nu_{n+1}'}{(n+1)!} u^{n+1}\right)\left(\frac{n+1}{n}\right)\right]^{1/2}$$

と書ける．ここで $\nu_{n-1}' = \mu_{n-1}'$ かつ $\nu_{n+1}' = \mu_{n+1}'$ だから，この不等式の右辺は $n \longrightarrow \infty$（n は奇数）のとき 0 となり，(5.5.2) の剰余項は 0 になる．したがって

(5.5.6)
$$\varphi(t+u) = 1 + \sum_{r=1}^{\infty} \frac{u^r}{r!} \varphi^{(r)}(t)$$

となり，級数は少なくとも，$|u| < c$ では収束する．$t = 0$ と置き，(5.1.5) を用いると

(5.5.7)
$$\varphi(u) = 1 + \sum_{r=1}^{\infty} \frac{\mu_r'}{r!} (\mathrm{i}u)^r$$

となり，これは $|u| < c$ での $\varphi(u)$ がモーメント μ_r' から一意的に定まることを意味する．解析接続により，公式 (5.5.7) はすべての u の値に対して成り立つ．なぜならば，$u = \pm \frac{1}{2} c$ で $\varphi(u)$ のすべての導関数が存在し，(5.5.7) より決定できる．したがって $|u| < c$ に対して，$\varphi\left(\frac{1}{2}c + u\right)$ と $\varphi\left(-\frac{1}{2}c + u\right)$ は (5.5.6) の級数の t を $\frac{1}{2}c$ と $-\frac{1}{2}c$ で置き換えることにより一意に定まる．これは，(5.5.7) が $|u| < \frac{3}{2}c$ で成り立つことに等しい．この方法を逐次繰り返すと，(5.5.7) が u のすべての値で成り立つことがいえる．このように，この定理の仮定の下での特性関数は，モーメント列により一意に定まり，逆に特性関数は **5.5.1** で述べたように，c.d.f. $F(x)$ を一意に定める．

次に示す **5.5.1** の系は，分布のモーメント列が与えられたとき，有限区間上で定義された分布関数を定めるのに役立つ．

5.5.1 a　x が有界確率変数であれば，その c.d.f. $F(x)$ はモーメント μ_r' $(r = 0, 1, 2, \cdots)$ により一意に定まる．

x が有界であれば，$F(a) = 0$，$F(b) = 1$，$a < b$ なる有限の数 a，b が存在する．M で $|a|$ と $|b|$ の大きい方を表わすと

(5.5.8)
$$|\mu_r'| \leqslant \int_a^b |x|^r\, dF(x) = \nu_r' \leqslant M^r,$$

(5.5.9)
$$\left|\sum_{r=0}^{\infty} \frac{\mu_r' c^r}{r!}\right| \leqslant \sum_{r=0}^{\infty} \frac{|\mu_r'| c^r}{r!} \leqslant \sum_{r=0}^{\infty} \frac{\nu_r' c^r}{r!} \leqslant \sum_{r=0}^{\infty} \frac{|Mc|^r}{r!} = e^{|Mc|}$$

となり，これは c のすべての値に対して有限である．したがって **5.5.1** の十分条件が満たされているから，**5.5.1 a** が証明された．

モーメント $\mu'_{r_1,\cdots,r_k}$ から k 次元確率変数 $(x_1, \cdots, x_k)$ の c.d.f. を一意に定めるための **5.5.1** と **5.5.1 a** の拡張は同様にして求まる．公式と証明は読者に残しておこう．

(b)　モーメントによる c.d.f. 列の極限の決定

$(x_1, x_2, \cdots)$ は確率変数列で，各成分のモーメント列が与えられているとする．このとき $(x_1, x_2, \cdots)$ がある確率変数 x に法則収束するかどうかという疑問が起るであろう．Kendall と Rao (1950) による簡単な定理があり，$(x_1, x_2, \cdots)$ からの部分列が c.d.f. へ法則収束する問題を 2 次モーメントだけを用いて述べている．それは次のようになる．

5.5.2　$\mu'_2(x_n)$ を確率変数列 $(x_1, x_2, \cdots)$ における x_n の 2 次モーメントとする．$n = 1, 2, \cdots$ に対して

(5.5.10)
$$\mu'_2(x_n) < \mathrm{K} < +\infty$$

であれば，法則収束する $(x_1, x_2, \cdots)$ の部分列が存在する．

この証明は，$F_n(x)$ を x_n の c.d.f. とし，任意の $x_0 > 0$ に対して

(5.5.11)
$$\mathrm{K} > \mu'_2(x_n) = \int_{-\infty}^{\infty} x^2\, dF_n(x) \geqslant x_0^2 \int_{-\infty}^{-x_0} dF_n(x) + x_0^2 \int_{x_0}^{\infty} dF_n(x)$$

が成り立つので，

(5.5.12)
$$\frac{\mathrm{K}}{x_0^2} > F_n(-x_0) + 1 - F_n(x_0), \qquad n = 1, 2, \cdots$$

と書ける．与えられた $\varepsilon > 0$ に対して $x > x_0$ で，すべての n に対して，$1 - [F_n(x) - F_n(-x)] < \varepsilon$ となる $x_0 > 0$ を選ぶことができる．**5.4.1** の証明で指摘したように，すべての連続点で，非減少関数 $G(x)$ に c.d.f. が収束するような，$(x_1, x_2, \cdots)$ の部分列を得ることができる．よって，$x > x_0$ に対して $1 - [G(x) - G(-x)] < \varepsilon$ となり，$G(-\infty) = 0$, $G(+\infty) = 1$ が成り立つ．それゆえ $G(x)$ は c.d.f. になる．これで確率変数 x へ法則収束する $(x_1, x_2, \cdots)$ の部分列に関する証明を終える．

さて，確率変数列 $(x_1, x_2, \cdots)$ の各成分 x_n に対して完備なモーメント列 $\mu'_r(x_n) = \mu'_{r,n}$, $r = 0, 1, 2, \cdots$ を持つと仮定し，かつ $\mu'_{r,n}$ はすべて有限で，$\lim_{n\to\infty} \mu'_{r,n} = \mu'_r$, $r = 0, 1, 2, \cdots$ であり，極限列 μ'_r, $r = 0, 1, 2, \cdots$ は c.d.f. $F(x)$ を一意に決定すると仮定する．確率変数列 $(x_1, x_2, \cdots)$ が $F(x)$ に等しい c.d.f. を持つ確率変数 x に法則収束するには，どのような条件が必要となるであろうか？

この疑問は Kendall と Rao (1950) により次のように解決されている．

5.5.3　$(x_1, x_2, \cdots)$ を確率変数列とする．x_n の r 次のモーメントを $\mu'_{r,n}$ とし，すべての n と r に対して有限とする．また $\lim_{n\to\infty} \mu'_{r,n} = \mu'_r$ とする．ただし μ'_r はすべての r に対して有限である．このとき，$(x_1, x_2, \cdots)$ が $F(x)$ に法則収束するならば，$\mu'_0, \mu'_1, \cdots$ は $F(x)$ のモーメント列である．逆に，このモーメント列が c.d.f. $F(x)$ を一意に決定すれば，それは $(x_1, x_2, \cdots)$ の極限 c.d.f. である．

5.5.3 の証明には，まず $(x_1, x_2, \cdots)$ が c.d.f. $F(x)$ に法則収束すると仮定して，$\mu'_r = \lim_{n\to\infty} \mu'_{r,n}$, $r = 1, 2, \cdots$ が $F(x)$ のモーメント列であることを示す．すなわち，$F_n(x)$ が x_n, $n = 1, 2, \cdots$ の c.d.f. であるとき

(5.5.13)
$$\lim_{n\to\infty} \left| \int_{-\infty}^{\infty} x^r\, dF_n(x) - \int_{-\infty}^{\infty} x^r\, dF(x) \right| = 0, \qquad r = 1, 2, \cdots$$

を示せば良い．

任意の $K > 0$ に対して

$$\left| \int_{-\infty}^{\infty} x^r\, dF_n(x) - \int_{-\infty}^{\infty} x^r\, dF(x) \right| \leqslant A_1 + A_2 + A_3$$

が成り立つ．ただし

(5.5.14)
$$A_1 = \left| \int_{-K}^{K} x^r\, dF_n(x) - \int_{-K}^{K} x^r\, dF(x) \right|$$
$$A_2 = \left| \int_{E_K} x^r\, dF_n(x) \right|, \qquad A_3 = \left| \int_{E_K} x^r\, dF(x) \right|$$

となり，E_K は $|x| > K$ なる x の値の集合である．これはシュワルツの不等式より

(5.5.15)
$$A_2^2 \leqslant \int_{E_K} x^{2r}\, dF_n(x) \cdot \int_{E_K} dF_n(x)$$

となり，両方の積分とも非負である．$n \longrightarrow \infty$ のとき，$\mu'_{2r,n} \longrightarrow \mu'_{2r}$（有限）だから，すべての n と K に対して最初の項の積分に対する上界 $M_r^2 > 0$ が存在する．また右辺の第 2 項の積分において，$n \longrightarrow \infty$ のとき $F_n(x) \longrightarrow F(x)$ となるので，A_2 は K を十分に大きくとると，すべての n に対していくらでも小さくできる．

μ'_r は有限なので，A_3 は K を十分に大きくとることにより任意に小さくできる．

$n \longrightarrow \infty$ のとき $F_n(x) \longrightarrow F(x)$, $\mu'_{r,n}$ と μ'_r はともに有限なので，任意の固定した K に対して n を十分大きくとれば，A_1 は任意に小さくできる．

よって，(5.5.13) の結論を得る．ゆえに，$\lim_{n\to\infty} \mu'_{r,n} = \mu'_r$, $r = 1, 2, \cdots$ で，$\mu'_1, \mu'_2, \cdots$ は $F(x)$ のモーメント列となり，$F(x)$ は $(x_1, x_2, \cdots)$ の c.d.f. 列の極限である．

さて逆を考えよう．$\mu_1', \mu_2', \cdots$ が c.d.f. $F(x)$ を一意に定めると仮定した場合，$F_1(x), F_2(x), \cdots$ は $F(x)$ という c.d.f. に収束することを示さねばならない．**5.5.2** から $F_1(x), F_2(x), \cdots$ のすべての収束部分列が，ある c.d.f. に収束することがわかり，また，いまの定理の最初の部分の議論より，これらの部分列に対する極限 c.d.f. は同じモーメント列，すなわち $\mu_1', \mu_2', \cdots$ を持たねばならない．しかしこのモーメント列は c.d.f. を一意に決定すると仮定されている．したがって，この極限 c.d.f. は $F(x)$ にすべて等しい．すなわち c.d.f. はモーメント列 $\mu_1', \mu_2', \cdots$ を持つ．

最後に，すべての n と r に対して $\mu_{r,n}'$ は有限である，という条件を次のような条件に置き換えても **5.5.3** が成り立つことを注意しておこう．これは r に依存するある整数 n^* より大きいすべての n と r に対して，$\mu_{r,n}'$ が有限であるということである．

問　　題

5.1　確率変数 x が縦軸に対称な p.d.f.（または p.f.）を持てば，x の特性関数 $\varphi(t)$ は実数値のみをとることを示せ．

5.2　確率変数 x が特性関数

$$\varphi(t) = \frac{1}{1+t^2}$$

を持てば，x の p.d.f. は R_1 における x の任意の値に対して，$\frac{1}{2}e^{-|x|}$ であることを示せ．

5.3　確率変数 x が特性関数

$$\frac{e^{it}(1-e^{nit})}{n(1-e^{it})}$$

を持てば，x は p.f. $p(x) = 1/n$, $x = 1, \cdots, n$ を持つ離散型確率変数であることを示せ．

5.4　$e^{-\frac{1}{2}t^2}$ を確率変数 x の特性関数としたとき，x の p.d.f. が $(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}x^2}$ であることを示せ．

5.5　**5.2.1** を証明せよ．

5.6　$x_1, \cdots, x_k$ が，c.d.f. がすべて $F(x)$ である独立な確率変数であれば，$x_1 + \cdots + x_k$ の特性関数は $[\varphi(t)]^k$ であることを示せ．ただし $\varphi(t)$ は c.d.f. $F(x)$ を持つ確率変数 x の特性関数である．

5.7　$-\infty < x < +\infty$ に対して，x の p.d.f. が

$$f(x) = \frac{k}{\pi[k^2 + (x-\mu)^2]}$$

であるとき，確率変数 x は **Cauchy** (1853) 分布を持つという．ただし k と μ は実定数で，$k > 0$ である．x の特性関数は

$$\varphi(t) = e^{i\mu t - k|t|}$$

となることを示せ．したがって，$x_1, \cdots, x_n$ がそれぞれ，このコーシー分布を持つ独立確率変数であれば，確率変数 $(x_1 + \cdots + x_n)/n$ もまた同じコーシー分布を持つ．

5.8　$x_1, \cdots, x_n$ が R_1 で同じ p.d.f.，すなわち $\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ を持つ独立確率変数であれば，確率変数 $\frac{1}{\sqrt{n}}(x_1 + \cdots + x_n)$ もまた同じ p.d.f. を持つことを特性関数を用いて示せ．

5.9　サイコロを 1 の目がでるまで投げた回数を確率変数 x で表わそう．x の確率母関数 $\theta(t)$ を求めよ．また $\theta(t)$ から x の平均と分散を求めよ．

5.10　“正しい” コインを k 回表が出るまで投げ，投げた回数を確率変数 x で表わす．$F(x;k)$ が x の c.d.f. であるとき，$F(x;k)$ が k に対して再性的であることを示せ．

5.11　x が，13 枚のトランプの中にはいっているスペードの枚数を表わす確率変数であれば，x の確率母関数 $\theta(t)$ が $(1+tv)^{13}(1+v)^{39}\Big/\binom{52}{13}$ における v^{13} の係数であることを示せ．また $\theta(t)$ から x の平均と分散を求めよ．

5.12　確率変数 x が

$$\mu_r' = \frac{k}{k+r}, \qquad r = 1, 2, \cdots$$

を持っている．ただし $k > 0$．x の p.d.f. が

$$f(x) = \begin{cases} kx^{k-1}, & 0 \leqslant x \leqslant 1 \\ 0, & \text{その他} \end{cases}$$

で与えられることと，x の分布が一意であることを示せ．

5.13　x が r 次のモーメント

$$\mu_r' = \frac{(k+r)!}{k!}, \qquad k \text{ は正の整数}$$

を持つ確率変数であれば，p.d.f. が

$$f(x) = \begin{cases} 0, & x \leqslant 0 \\ \dfrac{x^k}{k!}e^{-x}, & x > 0 \end{cases}$$

となることと，x の分布が一意であることを示せ．

5.14　(x_1, x_2) が，$r, s = 1, 2, \cdots$ に対して $\mathscr{E}(x_1^r x_2^s)$, $\mathscr{E}(x_1^r)$, $\mathscr{E}(x_2^s)$ のすべてが存在するような，有界な 2 次元確率変数で，$\mathscr{E}(x_1^r x_2^s) = \mathscr{E}(x_1^r)\cdot\mathscr{E}(x_2^s)$ が成り立てば，x_1 と x_2 が独立であることを示せ．

5.15　x を，n 個の “ひずみのない” サイコロを同時に投げたときの目の数の和を表わす確率変数とする．x のモーメント関数 $\psi(t)$ が $(e^t/6)^n(1-e^t)^{-n}(1-e^{6t})^n$ であることと，x の分布が n に関して再性的であることを示せ．

5.16 確率変数列 $(x_1, x_2, \cdots)$ で，x_n の p.d.f. は $(0,1)$ 上で $n(1-x)^{n-1}$ を持ち，その他で 0 となる．$y_n = nx_n$ のとき列 $(y_1, y_2, \cdots)$ が $(0,\infty)$ 上で e^{-y}，その他で 0 を持つ確率変数 y に法則収束することを特性関数を用いて示せ．

5.17 確率変数列 $(x_1, x_2, \cdots)$ で，x_n は特性関数

$$\varphi_n(t) = \frac{\sin nt}{nt}$$

を持つ．x_n の p.d.f. は $(-n, +n)$ 上で $1/2n$，その他で 0 を持つことを示せ．したがって，特性関数列が極限 $\varphi(t)$ に収束しても，c.d.f. の列は c.d.f. に収束しない．5.4.1 のどの条件がさまたげているのだろうか？ [Cramér (1946) を参照せよ]．

5.18 $(x_1, \cdots, x_k)$ は独立な離散型確率変数で，x_i の p.f. が

$$p_i(x) = \frac{\mu_i^x e^{-\mu_i}}{x!}, \qquad x = 0, 1, 2, \cdots, \qquad \mu_i > 0$$

であり，z が確率変数 $x_1 + \cdots + x_k$ であるとき，z の p.f. が

$$p(z) = \frac{\left(\sum_1^k \mu_i\right)^z e^{-\sum_1^k \mu_i}}{z!}, \qquad z = 0, 1, 2, \cdots$$

で与えられることを特性関数を用いて示せ．

5.19 $(x_1, x_2, \cdots)$ を，x_n の r 次モーメントが

$$\mu'_{r,n} = \frac{r!n^r(n-r-1)!}{(n-1)!},$$

$n > r = 1, 2, \cdots$ で与えられる非負の確率変数列とすれば，確率変数列は p.d.f.

$$f(x) = \begin{cases} 0, & x \leqslant 0 \\ e^{-x}, & x > 0 \end{cases}$$

を持つ確率変数 x に法則収束することを示せ．またこの極限分布はそのモーメントにより一意に定まることを示せ．

5.20 世代の移り変わりでの出生死亡，または原子核分裂過程での中性子のように，個体が次の個体を生成することができる（確率）分枝過程を考えてみよう．x_1 を初期（第0世代）の個体が，x_2 を第 1 世代が生成した個体の数を表わす確率変数とする．一般に x_{n+1} を第 n 世代が生成した個体の数を表わす確率変数とする．$x_1, x_2, \cdots$ の p.f. はすべて同一とする．すなわち $p(x)$，$x = 0, 1, 2, \cdots$ を持つ．y_n を第 n 世代で生成された個体の数の合計を表わす確率変数，$\theta_n(t)$ を y_n の確率母関数とする [注 $\theta_0(t) = t$，$\theta_1(t) = \sum_{x=0}^{\infty} t^x p(x)$]．確率母関数はすべて収束すると仮定すれば

$$\theta_{n+1}(t) = \theta_1[\theta_n(t)]$$

となることを示せ．

y_1 の平均と分散を μ と σ^2 で表わせば

$$\mathscr{E}(y_n) = \mu^n$$

$$\sigma^2(y_n) = n\sigma^2, \quad \mu = 1$$

$$= \sigma^2\mu^{n-1}(\mu^n - 1)/(\mu - 1), \qquad \text{その他}$$

となることを示せ．[Harris (1948)]．

5.21（続き） 確率変数 $x_1, x_2, \cdots, x_n$ がおのおの異なった p.f. すなわち $p_1(x), p_2(x), \cdots, p_n(x)$ と，おのおのに対応する確率母関数 $\theta_1^*(t), \theta_2^*(t), \cdots$ を持つとする．y_n の確率母関数 $\theta_n(t)$ は

$$\theta_n(t) = \theta_1^*[\theta_2^*\{\cdots\theta_n^*(t)\cdots\}]$$

で与えられることを示せ．したがって

$$\mathscr{E}(y_n) = \mu_1\mu_2\cdots\mu_n$$

となる．ただし $\mu_1, \cdots, \mu_n$ はそれぞれ $x_1, \cdots, x_n$ の平均である．

5.22 (5.4.4) を求めよ．

5.23 x が p.f. $p(x)$，c.d.f. $F(x)$ を持つ確率変数で，その標本空間として非負の整数の集合を持つとする．$F(x)$ に関する母関数 $\theta^*(t) = \sum_{x=0}^{\infty} F(x)t^x$ とする．$\theta(t)$ が x の確率母関数であれば，$|t| < 1$ に対して

$$\theta^*(t) = \theta(t)/(1-t)$$

となることを示せ．

5.24 $x_1, \cdots, x_k$ は標本空間が非負の整数で，確率母関数 $\theta_1(t), \cdots, \theta_k(t)$ を持つ独立な確率変数とする．$x_1 + \cdots + x_k$ の確率母関数が $\theta_1(t)\cdots\theta_k(t)$ となることを示せ．

第6章　離散型分布

本章では，数理統計学で取り扱われている重要な離散型確率分布をいくつか紹介する．それは分布そのものに関する基本的な情報を引き出すばかりでなく，第 1 章から第 5 章までに現われた概念，原理，方法などを，例題を用いて詳しく説明することにもなる．本章で議論された分布およびその主な性質は，後にいろいろな点で役立つであろう．さらに，前に紹介された一般的な原理や方法に関係なく，個々の分布の研究が特殊な方法や工夫を引き出してくれる場合もあるだろう．

6.1　超幾何分布

(a)　1 変数の場合

2.3(a) 節の終わりで用いられた確率関数 $p(x)$ は超幾何分布の簡単な場合である．一般的には次のようになる．Π を，その要素が C か $\overline{C}$ のいずれかに属する要素の集まりとし，Np を C に，Nq を $\overline{C}$ に属している要素の個数とする．ただし，$p+q=1$．いま，Π から $n(\leqslant N)$ 個の要素を取り出したとき，x を C に属している要素の数を表わす確率変数とする．x の分布関数を求めよう．Π の要素を n 個取り出す方法は $\binom{N}{n}$ 通りある．そのうち，x 個が C に含まれるのは，$\binom{Np}{x}\binom{Nq}{n-x}$ 通りである．

この場合，第 1 章で議論された基本標本空間 R は $\binom{N}{n}$ 個の標本点からなり，各標本点は Π から取り出された n 個の要素の集合である．$x \leqslant x'$ なる事象は，C に含まれている数が x' 以下である R の標本点全体からなっている．

R の各標本点に等確率を与えると，x の p.f. は次のようになる．

(6.1.1)
$$p(x)=\frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}.$$

ただし，R_1 における x の質点（x の標本空間）は $0 \leqslant x \leqslant Np$，$0 \leqslant n-x \leqslant Nq$ なる両不等式をともに満足する整数である．

これが**超幾何分布関数**である．この p.f. を持つ分布を**超幾何分布** $H(N,n\,;\,p)$ と書く．

x のとりうる値についての $p(x)$ の和が 1 になるのを見るには，恒等式
$$(u+v)^{A+B} \equiv (u+v)^A(u+v)^B$$
を考える．A，B は正の整数である．この両辺を展開すると
$$\sum_{r=0}^{A+B}\binom{A+B}{r}u^r v^{A+B-r} \equiv \left[\sum_{s=0}^{A}\binom{A}{s}u^s v^{A-s}\right]\left[\sum_{r=s}^{B+s}\binom{B}{r-s}u^{r-s}v^{B-r+s}\right]$$
が成り立つ．これは u，v に関する恒等式なので，両辺の $u^r v^{A+B-r}$ の係数は等しくなければならない．ゆえに

(6.1.2)
$$\binom{A+B}{r}=\sum_s \binom{A}{s}\binom{B}{r-s}.$$

ここで，$\sum_s$ は $0 \leqslant s \leqslant A$，$0 \leqslant r-s \leqslant B$ をともに満たす整数についての和を表わす．

(6.1.2) より

(6.1.2 a)
$$\sum_x \binom{Np}{x}\binom{Nq}{n-x}=\binom{N}{n}$$

は明らかである．ゆえに，(6.1.1) の x の標本空間にわたる $p(x)$ の和は 1 である．

この超幾何分布について，特性関数はモーメントの計算に対して実質的には，あまり役に立たない．しかし，階乗モーメント $\mu'_{[r]}$ は簡単に求まり，次のように有用である．

まず，

(6.1.3)
$$\sum_x x^{[r]}\binom{Np}{x}\binom{Nq}{n-x}=(Np)^{[r]}\sum_x \binom{Np-r}{x-r}\binom{Nq}{n-x}$$

が成立する．ここで，x は (6.1.2 a) とまったく同じ範囲を動く．(6.1.2) より，右辺の和は $\binom{N-r}{n-r}$．(6.1.3) を $\binom{N}{n}$ で割ると，左辺は $\mu'_{[r]}$ となり，その値は

(6.1.4)
$$\mu'_{[r]}=\frac{(Np)^{[r]}\binom{N-r}{n-r}}{\binom{N}{n}}$$

となる．特に

$$\mu'_{[1]} = np, \qquad \mu'_{[2]} = \frac{np(n-1)(Np-1)}{(N-1)}$$

であり，平均と分散は

$$(6.1.5) \qquad \mu(x) = np, \qquad \sigma^2(x) = \left(\frac{N-n}{N-1}\right)npq.$$

超幾何分布は初等確率論では重要な分布である．それは，カード遊びや，不良品を含んでいる大量生産物のロットおよび有限母集団からサンプルを取り出す場合の，確率の計算につながるからである．

超幾何分布 (6.1.1) の p.f. と c.d.f. は Lieberman と Owen (1961) によって幅広く表にまとめられている．

(b) k 変数の場合

確率分布 (6.1.1) は k 次元確率変数に一般化できる． 2.5 (a) 節の終わりの例題は 2 次元超幾何分布の特殊な場合である．一般的には，次のようになる．Π を N 個の要素の集まりとし，Np_i 個の要素が族 C_i $(i=1,\cdots,k+1)$ に属するものとする．ただし，$p_1+\cdots+p_{k+1}=1$ である． したがって， $C_1,\cdots,C_{k+1}$ は互いに素である． すなわち， Π の要素はこのどれか 1 つ， しかもただ 1 つの族に属する． さて， Π から n 個の要素の組を取り出す． $x_1,\cdots,x_{k+1}$ をそれぞれ $C_1,\cdots,C_{k+1}$ に属す要素の数を表わす確率変数とする．これらの確率変数は，その和が n であるから 1 次従属である．したがって，$x_1,\cdots,x_k$ を k 個の 1 次独立な確率変数とみなすことができ，x_{k+1} を $n-x_1-\cdots-x_k$ と表わせる．

Π からつくられる n 個の要素からなる可能な標本点が $\binom{N}{n}$ 通りある．すなわち，基本標本空間 R には $\binom{N}{n}$ 個の点がある．この標本空間の中には，x_1 個が C_1 に，$\cdots$，x_{k+1} 個が C_{k+1} にはいっている標本点がちょうど $\binom{Np_1}{x_1}\cdots\binom{Np_{k+1}}{x_{k+1}}$ 個ある． $\binom{N}{n}$ 個の点に等確率を与えると，$(x_1,\cdots,x_k)$ の p.f. は

$$(6.1.6) \qquad p(x_1,\cdots,x_k) = \frac{\binom{Np_1}{x_1}\cdots\binom{Np_{k+1}}{x_{k+1}}}{\binom{N}{n}}$$

となる．ただし，$x_{k+1}=n-x_1-\cdots-x_k$ である．これが，k 次元あるいは k 変数超幾何分布関数である．この分布の質点（標本点）は，$0 \leqslant x_i \leqslant Np_i$, $i=1,\cdots,k+1$ を満たす単体 $x_i \geqslant 0$, $i=1,\cdots,k$, $\sum_{i=1}^{k} x_i \leqslant n$ の 0 または正の整数を持つ R_k の点からなる．(6.1.6) を p.f. として持つ確率分布を k 変数超幾何分布 $H(N,n;p_1,\cdots,p_k)$ という．

(6.1.6) の平均，分散，共分散および高次のモーメントは階乗モーメントから求められる．階乗モーメント $\mathscr{E}(x_1^{[r_1]}\cdots x_k^{[r_k]})$ を $\mu'_{[r_1]\cdots[r_k]}$ で表わそう．(6.1.4) を導いた方法を拡張すると

$$(6.1.7) \qquad \mu'_{[r_1]\cdots[r_k]} = \frac{(Np_1)^{[r_1]}\cdots(Np_k)^{[r_k]}\binom{N-r_1-\cdots-r_k}{n-r_1-\cdots-r_k}}{\binom{N}{n}}$$

となる．ただし，$0 \leqslant r_i \leqslant Np_i$, $i=1,\cdots,k$, $r_1+\cdots+r_k \leqslant n$ である．(6.1.7) より

$$(6.1.8) \qquad \mu(x_i) = np_i, \qquad \sigma^2(x_i) = np_i(1-p_i)\left(\frac{N-n}{N-1}\right)$$

$$\mathrm{cov}\,(x_i,x_j) = -np_ip_j\left(\frac{N-n}{N-1}\right)$$

が確かめられる．

6.2 2 項分布

“操作” あるいは “試行” が行なわれたとき， 2 つの可能な結果， C と $\overline{C}$ のうちどちらか 1 つが現われるとする． たとえば，サイコロを投げたとき，1 の目が出た場合を C とし，その他の目が出た場合を $\overline{C}$ とする． p を C の起る確率， q を $\overline{C}$ の起る確率とすれば，$p+q=1$ となる．

“試行” が n 回繰り返されると，C と $\overline{C}$ からなる n 個の文字の列が得られる．2^n 個の可能な列がある． この列が基本標本空間 R をつくり，標本点は 2^n 個の可能な列からなっている． x を，その列の中で C の数が x 個であることを示す確率変数とする． n 回の “試行” が統計的に独立であると仮定すれば，$x=x'$ となる 1 つの列に割り当てられる確率は

$$(6.2.1) \qquad p^{x'}q^{n-x'}$$

である．これは列の C, $\overline{C}$ をそれぞれ p, q に置き換えて， 掛け合わせたものである．$x=x'$ となる列は $\binom{n}{x'}$ 通りあるので

(6.2.2)　$P(x = x') = \binom{n}{x'} p^{x'} q^{n-x'}, \qquad x' = 0, 1, \cdots, n.$

ゆえに，x は p.f. $p(x)$ が (6.2.2) の右辺（ダッシュを落す）で与えられる離散型確率変数である．$p(x)$ は $(p+q)^n$ を展開したときの一般項なので，(6.2.2) の右辺の x に関する和，すなわち，その確率変数 x の標本空間にわたる和は $(p+q)^n$ である．もちろん，$p+q=1$ であるから，この和も 1 である．p.f. (6.2.2) を持つ分布を **2 項分布**といい，$Bi(n, p)$ で表わす．(2 項分布に $Bi(a, b)$ を用い，第 7 章で紹介されるベータ分布に対しては記号 $Be(a, b)$ を用いる．) 2 つの可能な結果のうち 1 つが起る独立な試行列で，各試行とも，その結果が一定の確率で起っているものは，**ベルヌイ試行**と呼ばれている．これは J. Bernoulli (1713) によってはじめて研究された．

(6.2.2) の特性関数は

(6.2.3)　$\varphi(t) = \mathscr{E}(e^{itx}) = \sum_{x=0}^{n} \binom{n}{x} e^{itx} p^x q^{n-x} = (q + pe^{it})^n$

であるから，(5.1.5) よりモーメント μ'_r が計算される．特に，

(6.2.4)　$\mu'_1 = \frac{\varphi'(0)}{i} = np, \qquad \mu'_2 = -\frac{\varphi''(0)}{i^2} = np(q+np)$

となるから，(6.2.2) の平均と分散は

(6.2.5)　$\mu(x) = np, \qquad \sigma^2(x) = npq$

となる．

(6.2.3) で n をパラメータと考え，$\varphi(t)$ を $\varphi(t\,;n)$ と書けば，$\varphi(t\,;n)$ は (5.3.7) を満たす．よって，x_1，x_2 をそれぞれ 2 項分布 $Bi(n_1, p)$，$Bi(n_2, p)$ を持つ独立な確率変数とすれば，$x = x_1 + x_2$ は 2 項分布 $Bi(n_1 + n_2, p)$ に従う．ゆえに，

6.2.1　2 項分布 $Bi(n, p)$ は n に関して再生的である．

2 項分布 (6.2.2) は超幾何分布 (6.1.1) において，$N \longrightarrow \infty$ とした極限である．なぜなら，(6.1.1) は

(6.2.6)　$p(x) =$

$$\binom{n}{x} \frac{p\left(p-\frac{1}{N}\right)\cdots\left(p-\frac{x-1}{N}\right)q\left(q-\frac{1}{N}\right)\cdots\left(q-\frac{n-x-1}{N}\right)}{\left(1-\frac{1}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)} \xrightarrow[N\to\infty]{} \binom{n}{x} q^{n-x} p^x$$

となる．よって，n と x を固定すると次の結果を得る．

6.2.2　超幾何分布 $H(N, n\,;p)$ の p.f. において，$N \longrightarrow \infty$ のときの極限は 2 項分布 $Bi(n, p)$ の p.f. である．

2 項分布あるいは超幾何分布の質点の数は有限である．したがって，**6.2.2** より，$P_H(x \in E)$，$P_B(x \in E)$ をそれぞれ $H(N, n\,;p)$，$Bi(n\,;p)$ から計算された固定した事象 E の確率とすれば，$\lim_{N\to\infty} P_H(x \in E) = P_B(x \in E)$ となる．これを確率変数列で表わせば，次のようになる．x_N を超幾何分布 $H(N, n\,;p)$ を持つ確率変数とすれば，列 $(x_1, x_2, \cdots)$ は 2 項分布 $Bi(n, p)$ を持つ確率変数 x に分布収束する．

2 項分布は統計学における重要な分布の 1 つである．Army Ordnance Corps (1952), Harvard Computation Laboratory (1955), National Bureau of Standards (1949), Romig (1953) によって，p.f. (6.2.2) と c.d.f. 形がともに多岐にわたって表にされている．

6.3 多項分布

いま，"操作" あるいは "試行" によって，互いに排反な事象 $C_1, \cdots, C_{k+1}$ のうち，ただ 1 つだけが起り，その確率をそれぞれ $p_1, \cdots, p_{k+1}$ とし，すべて $p_i > 0$ で，$\sum_i p_i = 1$ とする．n 個の互いに独立な試行を行ない，$(x_1, \cdots, x_{k+1})$ を $(k+1)$ 次元確率変数で，$x_1, \cdots, x_{k+1}$ をそれぞれ $C_1, \cdots, C_{k+1}$ が起る試行の数を表わすものとする．$x_1 + \cdots + x_{k+1} = n$ だから，各成分 $x_1, \cdots, x_{k+1}$ は 1 次従属である．基本標本空間 R は $(k+1)^n$ 個の標本点からなり，各標本点は $C_1, \cdots, C_{k+1}$ から重複を許して，n 個を選択した列である．x_1 個が C_1 にはいり，…，x_{k+1} 個が C_{k+1} にはいっている標本点の確率は

(6.3.1)　$p_1^{x_1} \cdots p_{k+1}^{x_{k+1}}$

である．明らかに，このような標本点は R の中に

(6.3.2)　$\frac{n!}{x_1! \cdots x_{k+1}!}$

個ある．これらの各標本点の確率は (6.3.1) で与えられているので，確率変数 $(x_1, \cdots, x_k)$ の p.f. は

(6.3.3)　$p(x_1, \cdots, x_k) = \frac{n!}{x_1! \cdots x_{k+1}!} p_1^{x_1} \cdots p_{k+1}^{x_{k+1}}.$

ただし，$x_{k+1} = n - x_1 - \cdots - x_k$，$p_{k+1} = 1 - p_1 - \cdots - p_k$ である．$p(x_1, \cdots, x_k)$ は

多項式 $(p_1+\cdots+p_{k+1})^n$ の展開における一般項であることに注意しよう．p.f. (6.3.3) を持つ分布を k **変数多項分布**といい，$M(n\,;\,p_1,\cdots,p_k)$ で表わす．この分布の質点全体，すなわち $(x_1,\cdots,x_k)$ の標本空間は，単体 $x_i\geqslant 0,\ i=1,\cdots,k,\ \sum_{i=1}^{k}x_i\leqslant n$ に含まれる R_k の格子点からなる．これは単体に含まれる点で，その座標が 0 か正の整数値を持つものである．

多項分布の特性関数は

$$\varphi(t_1,\cdots,t_k)=\sum e^{it_1x_1+\cdots+it_kx_k}p(x_1,\cdots,x_k)$$
$$=\sum\frac{n!}{x_1!\cdots x_{k+1}!}(p_1e^{it_1})^{x_1}\cdots(p_ke^{it_k})^{x_k}(p_{k+1})^{x_{k+1}}.$$

$\sum$ は $(x_1,\cdots,x_k)$ の標本点全体にわたる和を表わす．しかし，これは多項式 $(p_1e^{it_1}+\cdots+p_ke^{it_k}+p_{k+1})^n$ の展開項のすべての和に他ならない．ゆえに

(6.3.4)　$$\varphi(t_1,\cdots,t_k)=(p_1e^{it_1}+\cdots+p_ke^{it_k}+p_{k+1})^n.$$

(5.2.3) を適用すると，$x_1,\cdots,x_k$ の結合モーメントがわかる．特に

(6.3.5)　$$\mu(x_i)=np_i,\qquad \sigma^2(x_i)=np_i(1-p_i)$$
$$\mathrm{cov}\,(x_i,x_j)=-np_ip_j$$

となる．

n を (6.3.4) でのパラメータとみなすと

6.3.1　多項分布 $M(n\,;\,p_1,\cdots,p_k)$ は n に関して再生的である．

6.4 ポアソン分布

2 項分布 (6.2.2) において x を固定する．$np=\mu$ として μ を固定して，$n\longrightarrow\infty$ とする．すなわち，$p\longrightarrow 0$ とする．このとき，(6.2.2) の極限は次の確率分布になる．

(6.4.1)　$$p(x)=\frac{\mu^xe^{-\mu}}{x!}.$$

これは発見者 Poisson (1837) の名にちなんで，**ポアソン分布**といわれている．(6.4.1) の p.f. を持つ分布を $Po(\mu)$ と書く．(確率の記号 $P(\mu)$ との混乱を避けるため $Po(\mu)$ というように "o" をつける.) (6.4.1) を得るのに n を限りなく大きくしたので，x は 0 または正の整数をとることができる．このように x の質点は $0,1,2,\cdots$ である．(6.4.1) を導くには，(6.2.2) の右辺を

(6.4.2)　$$\frac{\dfrac{n}{n}\left(1-\dfrac{1}{n}\right)\left(1-\dfrac{2}{n}\right)\cdots\left(1-\dfrac{x-1}{n}\right)}{x!}(np)^x\left(1-\frac{np}{n}\right)^{n-x}$$

と書き直す．$np=\mu$ として，$n\longrightarrow\infty$ すなわち $p\longrightarrow 0$ とすれば，(6.4.2) の極限が (6.4.1) の分布になる．(6.4.1) の，$x=0,1,2,\cdots$ にわたる確率の和が 1 になることは容易に確かめられる．

ポアソン分布の特性関数は

(6.4.3)　$$\varphi(t)=\sum_{x=0}^{\infty}\frac{e^{itx}\mu^xe^{-\mu}}{x!}=e^{-\mu(1-e^{it})}$$

となり

(6.4.4)　$$\mu(x)=\mu,\qquad \sigma^2(x)=\mu$$

となる．

μ を (6.4.3) の特性関数のパラメータと考えると，(5.3.7) を満たすから

6.4.1　ポアソン分布 $Po(\mu)$ は μ に関して再生的である．

注意　2 項分布は十分大きな N に対する超幾何分布への近似であり，ポアソン分布は十分大きな n，小さな p に対する 2 項分布への近似である．したがって，比較的簡単なポアソン分布は，特定の条件の下では超幾何分布の近似になる．簡単にいえば，これは (i) n が十分大きく，(ii) p を小さく ($np=\mu$ で一定)，かつ (iii) N が n に比べて相当に大きい場合である．この条件は次の適用例では満たされている．たとえば，N をロットの大きさ，n をそのロットから取り出された標本の大きさ，p を不良品の割合とした，大量生産物のロットから抽出する場合である．

2 項分布と同様に，ポアソン分布も確率および統計の応用に際しては大変重要な分布の 1 つである．工場での大量生産物の標本に含まれる不良品の分布ばかりでなく，培養中あるいは溶液中の単位体積当りのバクテリアの分布，単位時間内に電話がかかる回数，単位区画当りの標的に放たれた爆弾の数などに，ポアソン分布が表われる．いずれの場合でも，独立あるいはほとんど独立に近い "試行" の回数 n は十分大きく，どの試行においても "成功する" 確率 p は小さく，n 回の試行のうち x 回 "成功する" 確率が問題になる．

ポアソン分布の p.f. と c.d.f. は，Molina (1942) が表にしている．

6.5 離散的待ち時間分布

(a) 超幾何分布の場合

$p>0,\ q>0,\ p+q=1$ とする Π を，C が Np 個はいっていて，$\overline{C}$ が Nq 個はいっている要素の集合とする．いま，Π から逐次要素を取り出し，C の要素をちょうど k 個取り出すまで行なう．Π から，ちょうど x 個取り出したときに，C が k 個はいっている確率を求めよう．基本標本空間 R は，Π から N 個の要素すべてを取り出した C と $\overline{C}$ からなる順列全体である．したがって，R の標本点の数は

(6.5.1)
$$\binom{N}{Np}$$

である．各標本点に対して確率変数 x は，ちょうど k 個の C を引き出すため抽出された Π の要素の数である．$0\leqslant k-1\leqslant x'-1,\ 0\leqslant Np-k\leqslant N-x'$ となる x' に対して，$x=x'$ となる標本点の数は，組合せにより

(6.5.2)
$$\binom{x'-1}{k-1}\binom{N-x'}{Np-k}$$

となる．R の各標本点に等確率 $1\Big/\binom{N}{Np}$ を与えれば，

(6.5.3)
$$P(x=x')=\frac{\binom{x'-1}{k-1}\binom{N-x'}{Np-k}}{\binom{N}{Np}}$$

が（ダッシュを落して）確率変数 x の p.f. $p(x)$ になる．(6.5.3) の p.f. を持つ分布を**超幾何待ち時間分布**と呼ぶ．確率変数 x の特定の値 x' は，本質的には k 個の C を得るために待ち続けなければならない試行の数である．

(6.5.4)
$$p(x)=\binom{x-1}{k-1}\binom{N-x}{Np-k}\Big/\binom{N}{Np}$$

が p.f.，すなわち x が $k,k+1,\cdots,Nq+k,\ 1\leqslant k\leqslant Np$ を動くとき，$p(x)$ の和が 1 となることを確かめるには

(6.5.5)
$$\sum_{x=k}^{Nq+k}\binom{x-1}{k-1}\binom{N-x}{Np-k}=\binom{N}{Np}$$

を示せばよい．

$|t|<1$ に対して

(6.5.6)
$$(1-t)^{-k}=\sum_{x=k}^{\infty}\binom{x-1}{k-1}t^{x-k}$$
$$(1-t)^{-(Np-k+1)}=\sum_{y=-\infty}^{Nq+k}\binom{N-y}{Np-k}t^{Nq-y+k}$$

(6.5.7)
$$(1-t)^{-(Np+1)}=\sum_{z=0}^{\infty}\binom{Np+z}{Np}t^{z}.$$

(6.5.7) より，$(1-t)^{-(Np+1)}$ の展開項における t^{Nq} の係数は $\binom{N}{Np}$ である．(6.5.6) の両式を掛け合わせると

(6.5.8)
$$(1-t)^{-(Np+1)}=\sum_{x}\sum_{y}\binom{x-1}{k-1}\binom{N-y}{Np-k}t^{Nq+(x-y)}.$$

(6.5.8) において，$(1-t)^{-(Np+1)}$ を展開すると，t^{Nq} の係数は，$\binom{x-1}{k-1}\binom{N-y}{Np-k}$ を $x=y$ なる整数 (x,y) の可能な対全体で動かしたときの和である．すなわち

(6.5.9)
$$\sum_{x=k}^{Nq+k}\binom{x-1}{k-1}\binom{N-x}{Np-k}.$$

$(1-t)^{-(Np+1)}$ の展開で，t^{Nq} の係数は $\binom{N}{Np}$ となり，一方，(6.5.9) にもなる．ゆえに，(6.5.5) が成り立ち，(6.5.4) で定義された $p(x)$ は p.f. になる．

(6.5.4) の分布では，その特性関数はモーメントの計算には役に立たないが，6.1 節と同様にして，そのモーメントが求められる．この場合，$\mathscr{E}[(x-k)^{[r]}]$，すなわち k に関する x の r 次の階乗モーメントを評価するのが容易なので，これからモーメントが計算される．

(6.5.10)
$$\mathscr{E}[(x-k)^{[r]}]=\sum_{x=k}^{Nq+k}(x-k)^{[r]}\binom{x-1}{k-1}\binom{N-x}{Np-k}\Big/\binom{N}{Np}$$
$$=(k+r-1)^{[r]}\left\{\sum_{x=k+r}^{Nq+k}\binom{x-1}{k+r-1}\binom{N-x}{Np-k}\right\}\Big/\binom{N}{Np}.$$

(6.5.4) の右辺における x の標本空間にわたる和は $\binom{N}{Np}$ であるから，{ } 内では $\binom{N}{Np+r}$ となる．したがって

(6.5.11)
$$\mathscr{E}[(x-k)^{[r]}]=(k+r-1)^{[r]}\binom{N}{Np+r}\Big/\binom{N}{Np}.$$

$r=1,2$ とすれば，平均と分散は

(6.5.12)
$$\mu(x)=\frac{k(N+1)}{Np+1}$$

$$\sigma^2(x)=\frac{Nq(Nq-1)k(k+1)}{(Np+2)(Np+1)}+\frac{k(2k+1)(N+1)}{Np+1}-\frac{k^2(N+1)^2}{(Np+1)^2}-k(k+1). \tag{6.5.13}$$

(b) 2 項分布の場合

互いに独立な“試行”を連続的に行ない，その結果を C か $\overline{C}$ のいずれかとし，$P(C)=p$，$P(\overline{C})=q$，$p+q=1$ とする．標本空間 R は C と $\overline{C}$ からなる無限に続く列の集合である．各標本点に対して，x はちょうど k 個の C を得るまでになされた試行の数を表わす確率変数となる．

E を，最初の $(x-1)$ 回の試行で $(k-1)$ 個 C を得た R における事象，F を x 回目に C を得る事象とする．x 回の試行で，ちょうど k 個 C を得る事象を G とすれば，$G=E\cap F$ であり，$P(G)=P(E\cap F)=P(E)\cdot P(F|E)$ となる．しかし，$P(F|E)=p$ で

$$P(E)=\binom{x-1}{k-1}p^{k-1}q^{x-k}. \tag{6.5.14}$$

ゆえに，x の p.f. を $p(x)$ とすれば，$p(x)=P(G)$，すなわち

$$p(x)=\binom{x-1}{k-1}p^kq^{x-k},\qquad x=k,k+1,\cdots. \tag{6.5.15}$$

$(1-q)^{-k}$ を q についてべき級数展開すれば，$p(x)$ は $p^k(1-q)^{-k}$ を展開して得られる $(x-k+1)$ 番目の項であることがわかる．ゆえに，$p(x)$ の $x=k,k+1,\cdots$ にわたる和は $p^k(1-q)^{-k}=p^kp^{-k}=1$ となる．

(6.5.15) で定義される p.f. を **2 項待ち時間分布**という．p.f. $p(x)$ は簡単にいえば，k 個 C を得るまでに x 回独立な試行を行なわねばならない確率である．この分布は，**負の 2 項分布，パスカル分布**とも呼ばれている．

(6.5.15) の特性関数は

$$\varphi(t)=\sum_{x=k}^{\infty}\binom{x-1}{k-1}(pe^{it})^k(qe^{it})^{x-k}=(pe^{it})^k(1-qe^{it})^{-k}$$

だから

$$\varphi(t)=p^k(e^{-it}-q)^{-k}. \tag{6.5.16}$$

微分して

$$\mu(x)=\frac{k}{p},\qquad \sigma^2(x)=\frac{kq}{p^2} \tag{6.5.17}$$

が確かめられる．

この平均と分散は (6.5.12) と (6.5.13) で $N\longrightarrow\infty$ とした極限である．事実，2 項待ち時間分布 (6.5.15) の p.f. は超幾何待ち時間分布 (6.5.3) の p.f. で $N\longrightarrow\infty$ とした極限である．これを確かめるのは，練習問題としておこう．

この節では，もっとも簡単な離散待ち時間分布だけしか考えなかった．単に C のみを特定個数だけ得るために要する試行の回数の確率ばかりでなく，C と $\overline{C}$ がある特定個数に達するまでに要する試行の回数の確率も算出したい．また結果が 2 つ以上の試行についても，特定個数の結果を得るための試行の必要回数の確率も算出したい．この種の問題はここで議論したよりも，さらに複雑になる．このうちのいくつかは，Girshick, Mosteller と Savage(1946)，Haldane (1945)，Laplace (1814)，McCarthy (1947) らが計算している．

6.6 連の理論における分布

第 14 章で見るように，連の理論はノンパラメトリック統計的推測の問題で，大切な役割を果たしている．しかし，ここでは，この理論の主要な分布だけをまとめて紹介するにとどめる．以下の基本的な結果は Mood (1940) に依るところが大きい．

(a) 2 種類の要素からなる連

標本空間 R として，n_1 個の C と n_2 個の $\overline{C}$ の $\binom{n}{n_1}$ 個の順列からなる標本点全体を考える．ただし $n_1+n_2=n$．任意の標本点は C と $\overline{C}$ からなる順列で，C の連と $\overline{C}$ の連が交互に現われている．連に含まれている要素の数を**連の長さ**という．r_{1j} を長さ j の C の連の数，r_{2j} を長さ j の $\overline{C}$ の連の数とする．たとえば，順列が

$$CCC\overline{C}\overline{C}CC\overline{C}CC\overline{C}\overline{C}C\overline{C}$$

であれば，$n_1=8$，$n_2=6$，$r_{11}=1$，$r_{12}=2$，$r_{13}=1$，$r_{21}=2$，$r_{22}=2$，その他の r_{ij} は 0 である．

n_1，n_2，r_{1j}，r_{2j} の定義より，$\sum_j jr_{1j}=n_1$，$\sum_j jr_{2j}=n_2$ である．$r_1=\sum_j r_{1j}$，$r_2=\sum_j r_{2j}$ とすれば，r_1，r_2 はそれぞれ C，$\overline{C}$ の連の数を表わす．R の各標本点に特定の確

率を割り当てれば，確率変数 r_{1j}，r_{2j} を定義したことになり，それは R の任意の標本点に対して上述の値を持つ．R の各標本点に等確率を与えたときに，これらの確率変数の p.f. を見つけるのが問題である．各 r_{1j} の組に対して，r_1 個の C の連を並べる方法は

$$(6.6.1)\qquad \frac{r_1!}{r_{11}!\cdots r_{1n_1}!}$$

通りある．同様に，r_2 個の $\overline{C}$ の連の並べ方は

$$(6.6.2)\qquad \frac{r_2!}{r_{21}!\cdots r_{2n_2}!}$$

通りである．

r_1 と r_2 は高々 1 つしか違わない．そうでないならば，たとえば，C の連のうち少なくとも 2 つ以上の連が隣り合っていなければならない．しかし，これは連の定義に矛盾する．$r_1=r_2$ ならば，C の連の組と $\overline{C}$ の連の組はどちらかを先に並べて，以下交互に並べればよいから，その連としての並べ方は 2 通りある．

$\gamma(r_1,r_2)$ を，r_1 個の区別できないある対象と，r_2 個の区別できない別の対象とを並べ，同じ対象が隣り合わないように並べる方法の数とする．

$$(6.6.3)\qquad \gamma(r_1,r_2)=\begin{cases}0, & |r_1-r_2|>1\\ 1, & |r_1-r_2|=1\\ 2, & r_1=r_2\end{cases}$$

である．長さ $j\ (=1,\cdots,n_1)$ の C の r_{1j} 個の連と，長さ $j\ (=1,\cdots,n_2)$ の $\overline{C}$ の r_{2j} 個の連を並べる方法は全部で

$$(6.6.4)\qquad \frac{r_1!r_2!\gamma(r_1,r_2)}{r_{11}!\cdots r_{1n_1}!r_{21}!\cdots r_{2n_2}!}$$

通りになる．しかし，R の標本点には n_1 個の C と n_2 個の $\overline{C}$ からなる $\binom{n}{n_1}$ 個の並べ方がある．この $\binom{n}{n_1}$ 個の標本点に等確率を与えれば，(n_1+n_2) 次元確率変数 $(r_{ij};\ j=1,\cdots,n_i,\ i=1,2)$ は次の p.f. を持つ．

$$(6.6.5)\qquad p(\{r_{ij}\})=\frac{r_1!r_2!}{r_{11}!\cdots r_{1n_1}!r_{21}!\cdots r_{2n_2}!}\cdot\frac{n_1!n_2!}{n!}\cdot\gamma(r_1,r_2).$$

C の連，すなわち $r_{1j}\ (j=1,\cdots,n_1)$ だけに注目するならば，この p.f. としては (6.6.5)で，$(r_{11},\cdots,r_{1n_1})$ の周辺分布を求めればよい．すなわち，$p(\{r_{ij}\})$ を $r_{21},\cdots,r_{2n_2}$ に関して和をとる．これは (6.6.2) を $r_{21},\cdots,r_{2n_2}$ に関して，$\sum_j jr_{2j}=n_2$，$\sum_j r_{2j}=r_2$ なる範囲で加えることである．このためには，s に関する恒等式で，その 0 近傍で成り立つ次式を用いる．

$$(6.6.6)\qquad (s+s^2+\cdots)^{r_2}\equiv s^{r_2}(1-s)^{-r_2}\equiv s^{r_2}\sum_{t=0}^{\infty}\frac{(r_2+t-1)!}{(r_2-1)!t!}s^t.$$

(6.6.6) の左辺の s^{n_2} の係数は，(6.6.2) を $r_{21},\cdots,r_{2n_2}$ について $\sum_j jr_{2j}=n_2$，$\sum_j r_{2j}=r_2$ なる範囲で加えた和である．(6.6.6) の左辺から，s^{n_2} の係数は (6.6.6) の右辺の s^{n_2} の係数に等しく

$$(6.6.7)\qquad \frac{(n_2-1)!}{(r_2-1)!(n_2-r_2)!}$$

である．ゆえに，$(r_{1j};\ j=1,\cdots,n_1)$ と r_2 との p.f. は

$$(6.6.8)\qquad p(\{r_{1j}\},r_2)=\frac{r_1!}{r_{11}!\cdots r_{1n_1}!}\cdot\frac{(n_2-1)!}{(r_2-1)!(n_2-r_2)!}\cdot\frac{n_1!n_2!}{n!}\cdot\gamma(r_1,r_2).$$

r_{1j} だけの p.f. は r_2 について，(6.6.8) を加え合わせればよい．(6.6.3) を用いると

$$(6.6.9)\qquad \sum_{r_2=1}^{n_2}\frac{(n_2-1)!}{(r_2-1)!(n_2-r_2)!}\cdot\gamma(r_1,r_2)=\frac{(n_2+1)!}{r_1!(n_2-r_1+1)!}=\binom{n_2+1}{r_1}.$$

ゆえに，$(r_{1j};\ j=1,\cdots,n_1)$ の p.f. は

$$(6.6.10)\qquad p(\{r_{1j}\})=\frac{r_1!}{r_{11}!\cdots r_{1n_1}!}\cdot\binom{n_2+1}{r_1}\Big/\binom{n}{n_1}.$$

r_{2j} の分布も同様にして求められる．

r_1 と r_2 との分布も重要である．この分布は，(6.6.8) を r_{1j} について，$\sum_j jr_{1j}=n_1$，$\sum_j r_{1j}=r_1$ の範囲で加えて，(6.6.5) から (6.6.8) を導いたようにすると

$$(6.6.11)\qquad p(r_1,r_2)=\frac{\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_2-1}}{\binom{n}{n_1}}\gamma(r_1,r_2).$$

r_1，すなわち C の連の総数の p.f. は ((6.6.11) を r_2 に関して加え合わせると)

$$(6.6.12)\qquad p(r_1)=\frac{\binom{n_1-1}{r_1-1}\binom{n_2+1}{r_1}}{\binom{n}{n_1}}$$

である．

r_2 の p.f. も同様に求められる．

r_1 が標本空間を動くときの $p(r_1)$ の和が 1 であることを確かめるには

$$\sum_{i=0}^{B}\binom{A}{K+i}\binom{B}{i}=\binom{A+B}{K+B} \tag{6.6.13}$$

を用いる．これは恒等式

$$(1+s)^A\left(1+\frac{1}{s}\right)^B\equiv\frac{(1+s)^{A+B}}{s^B} \tag{6.6.14}$$

において，s^K の係数が等しいことから得られる．(6.6.13) より

$$\sum_{r_1=1}^{n_1}\binom{n_1-1}{r_1-1}\binom{n_2+1}{r_1}=\binom{n}{n_1}.$$

ゆえに，

$$\sum_{r_1=1}^{n_1}p(r_1)=1.$$

r_1 のモーメントを見つける一番よい方法は階乗モーメントからであろう．g 次の階乗モーメント $\mu'_{[g]}$ は

$$\mu'_{[g]}=\mathscr{E}(r_1^{[g]})=\frac{(n_2+1)^{[g]}}{\binom{n}{n_1}}\sum_{r_1=g}^{n_1}\binom{n_1-1}{r_1-1}\binom{n_2+1-g}{r_1-g}. \tag{6.6.15}$$

(6.6.13) より

$$\sum_{r_1=g}^{n_1}\binom{n_1-1}{r_1-1}\binom{n_2+1-g}{r_1-g}=\binom{n-g}{n_1-g}.$$

ゆえに，

$$\mu'_{[g]}=(n_2+1)^{[g]}\frac{\binom{n-g}{n_1-g}}{\binom{n}{n_1}}. \tag{6.6.16}$$

平均と分散は

$$\mu(r_1)=\frac{n_1(n_2+1)}{n},\qquad \sigma^2(r_1)=\frac{(n_2+1)^{[2]}(n_1)^{[2]}}{n(n)^{[2]}}. \tag{6.6.17}$$

r_2 の平均と分散についても，階乗モーメントに関する式が同様に成り立つ．

(6.6.11) に対しても，同様に (r_1-1) と (r_2-1) に関する一般の階乗モーメントが求められる．すなわち

$$\mathscr{E}[(r_1-1)^{[g_1]}(r_2-1)^{[g_2]}]=\frac{(n_1-1)^{[g_1]}(n_2-1)^{[g_2]}}{\binom{n}{n_1}}\binom{n-g_1-g_2}{n_1-g_2}. \tag{6.6.18}$$

特に，連の総数を表わす確率変数，すなわち $u=r_1+r_2$ は重要である．u の平均と分散は (6.6.17)，(6.6.18) から，(3.4.2)，(3.4.3) が適用できて

$$\mu(u)=\mu(r_1)+\mu(r_2)=\frac{2n_1n_2}{n}+1 \tag{6.6.19}$$

$$\sigma^2(u)=\sigma^2(r_1)+\sigma^2(r_2)+2\,\mathrm{cov}\,(r_1,r_2)=\frac{2n_1n_2(2n_1n_2-n)}{n^2(n-1)}.$$

事実，u の p.f. は (6.6.11) から計算される．それは $p(r_1,r_2)$ を r_1r_2 平面上の直線 $u=r_1+r_2$ に沿って加え合わせればよい．u が偶数のとき，$p(r_1,r_2)$ が 0 でない点はただ 1 点だけである．奇数のときには 2 点になる．よって，u の p.f. は

$$p(u)=\begin{cases}\dfrac{2\binom{n_1-1}{\frac{1}{2}u-1}\binom{n_2-1}{\frac{1}{2}u-1}}{\binom{n}{n_1}}, & u\text{ が偶数のとき}\\[3ex] \dfrac{\binom{n_1-1}{\frac{1}{2}u-\frac{1}{2}}\binom{n_2-1}{\frac{1}{2}u-\frac{3}{2}}+\binom{n_1-1}{\frac{1}{2}u-\frac{3}{2}}\binom{n_2-1}{\frac{1}{2}u-\frac{1}{2}}}{\binom{n}{n_1}}, & u\text{ が奇数のとき}\end{cases} \tag{6.6.20}$$

と書き表わされる．これは Stevens (1939) がはじめて得た．$p(u)$ に関する各種の表が Swed と Eisenhart (1943) によってつくられている．

いままでの議論では n_1 と n_2 は固定されていた．これを確率変数とみなせば，p.f. がそれぞれ (6.6.5)，(6.6.10)，(6.6.11)，(6.6.12) で与えられる分布は n_1 と n_2 を固定したときの条件つき分布になる．n_1 と n_2 の分布を $p^*(n_1,n_2)$ とすれば，この 4 つの分布にそれぞれ $p^*(n_1,n_2)$ を掛けたものが，r と n_1，n_2 との p.f. になる．ともかく r だけの p.f. を求めるには，その積分布を n_1 と n_2 の動く範囲にわたって加え合わせればよい．特に，結果が C か $\overline{C}$ のいずれかである独立試行の一組を考えると，n_1，n_2 は 1 次従属で，$n_1+n_2=n$ を満たし，n_1 は 2 項分布 $Bi(n,p)$ に従うであろう．ただし p は各試行で C が起る確率である．この場合

$$p^*(n_1,n_2)=\binom{n}{n_1}p^{n_1}q^{n_2}. \tag{6.6.21}$$

ただし，$n_2=n-n_1$．たとえば，r_1 の g 次階乗モーメントを見つけるには (6.6.16) の右辺と (6.6.21) を掛け，$n_1=0,1,\cdots,n$ にわたって加えればよい．もちろん，$n=1,\cdots,g-1$ に対するこの和の項は 0 になることがわかる．

(b) k 種類の要素からなる連

前述の結果は，数種類の要素から生成された連の場合にも直接拡張可能である．n_1 個の C_1, …, n_k 個の C_k からなる，全体で n 個の要素を考える．もちろん，$n_1+\cdots+n_k=n$. r_{ij} を長さ j の C_i の連の数を表わす確率変数とする．$r_i=\sum_j r_{ij}$ は C_i の連の総数である．Mood (1940) は $\{r_{ij}\}$ の p.f. を導いている．すなわち

(6.6.22)　$$p(\{r_{ij}\})=\frac{n_1!\cdots n_k!}{n!}\left[\prod_{i=1}^{k}\frac{r_i!}{r_{i1}!\cdots r_{in_i}!}\right]\cdot\gamma(r_1,\cdots,r_k).$$

ただし，$\gamma(r_1,\cdots,r_k)$ は，r_1 個のある対象，r_2 個の別の対象，… を並べる方法で，そのうち同じものが隣り合わない並べ方の数を表わす．関数 $\gamma(r_1,\cdots,r_k)$ は

(6.6.23)　$$(x_1+\cdots+x_k)^k(x_2+x_3+\cdots+x_k)^{r_1-1}\cdot(x_1+x_3+\cdots+x_k)^{r_2-1}\cdots(x_1+x_2+\cdots+x_{k-1})^{r_k-1}$$

を展開したときの $x_1^{r_1}\cdots x_k^{r_k}$ の係数である．(6.6.22) を導き出すのは (6.6.5) のときと同様の方法である．これは読者にゆだねる．(6.6.5) から (6.6.11) を導いたように，$(r_1,\cdots,r_k)$ の p.f. は

(6.6.24)　$$p(\{r_i\})=\frac{n_1!\cdots n_k!}{n!}\left[\prod_{i=1}^{k}\binom{n_i-1}{r_i-1}\right]\cdot\gamma(r_1,\cdots,r_k).$$

$k=2$ の場合と同様に，(6.6.22) と (6.6.24) からモーメントが計算される．

問　題

6.1　x が超幾何分布 $H(N,n;p)$ に従うとする．このとき

$$\mathscr{E}[(n-x)^{[r]}]=(Nq)^{[r]}\binom{N-r}{n-r}\Big/\binom{N}{n}$$

を示せ．

6.2　壺の中に $1,2,\cdots,m+n$ と番号のついている $(m+n)$ 枚の札がはいっている．壺からランダムに n 枚の札を取り出す．このとき，壺に残っているどの札の番号よりも大きい番号をつけた札が，x 枚取り出される確率は

$$\binom{m+n-x-1}{m-1}\Big/\binom{m+n}{m}$$

であることを示せ．また

$$\mathscr{E}[(m+n-x+r-1)^{[r]}]=(m+r-1)^{[r]}\binom{m+n+r}{m+r}\Big/\binom{m+n}{n}$$

を示せ．これから x の平均と分散を求めよ．

6.3　$(x_1,\cdots,x_k)$ を k 変数超幾何分布 $H(N,n;p_1,\cdots,p_k)$ に従う k 次元確率変数とする．周辺分布 $(x_1,\cdots,x_{k_1})$, $k_1<k$ は超幾何分布 $H(N,n;p_1,\cdots,p_{k_1})$ となることを示せ．

6.4 (続き)　条件つき確率変数 $x_k|x_1,\cdots,x_{k-1}$ は 1 次元超幾何分布

$$H\left(N(p_k+p_{k+1}),n-x_1-\cdots-x_{k-1};\frac{p_k}{p_k+p_{k+1}}\right)$$

に従うことを示せ．

6.5 (続き)　$x_1+\cdots+x_k$ は 1 次元超幾何分布 $H(N,n;p_1+\cdots+p_k)$ を持つことを示せ．

6.6　(6.1.6) の $p(x_1,\cdots,x_k)$ の全標本空間上での和が 1 であることを示せ．

6.7　**6.2.1** を証明せよ．

6.8　**6.3.1** を証明せよ．

6.9　**6.4.1** を証明せよ．

6.10　$np_1=\mu_1,\cdots,np_k=\mu_k$ となるように，$n\longrightarrow\infty$ したがって $p_1,\cdots,p_k$ をそれぞれ $\longrightarrow 0$ とすれば，(6.3.3) における $(x_1,\cdots,x_k)$ の p.f. は独立なポアソン分布 $Po(\mu_1)$, …, $Po(\mu_k)$ の p.f. の積になることを示せ．

6.11　(6.5.15) の 2 項待ち時間分布は k に関して再生的であることを示せ．

6.12　$(x_1,\cdots,x_k)$ が (k 変数) 多項分布 $M(n;p_1,\cdots,p_k)$ を持つとき，特性関数を用いて，$(x_1,\cdots,x_{k_1})$, $k_1<k$ の周辺分布が (k_1 変数) 多項分布 $M(n;p_1,\cdots,p_{k_1})$ に従うことを示せ．

6.13 (続き)　条件つき確率変数 $x_k|x_1,\cdots,x_{k-1}$ は 2 項分布

$$Bi\left(n-x_1-x_2-\cdots-x_{k-1};\frac{p_k}{p_k+p_{k+1}}\right)$$

に従うことを示せ．

6.14 (続き)　$x_1+\cdots+x_{k_1}$, $k_1\leqslant k$ は 2 項分布 $Bi(n;p_1+\cdots+p_{k_1})$ を持つことを示せ．

6.15　(6.5.3) の超幾何待ち時間分布で，$N\longrightarrow\infty$ とすれば，極限は (6.5.15) の 2 項待ち時間分布になることを示せ．

6.16　(6.6.12) を確かめよ．

6.17　(6.6.18) を確かめよ．

6.18　x がポアソン分布 $Po(\mu)$ に従い，条件つき確率変数 $y|x$ が 2 項分布 $Bi(x,p)$ を持つとき，y 自身はポアソン分布 $Po(\mu p)$ に従うことを示せ．

6.19　2 項待ち時間分布 (6.5.15) で $y=x-k$ とする．$kq\longrightarrow\mu$ (一定) で，$k\longrightarrow\infty$, $q\longrightarrow 0$ とすれば y はポアソン分布 $Po(\mu)$ に従うことを示せ．

6.20　x がポアソン分布 $Po(\mu)$ に従うとき，x の c.d.f. は

$$\frac{1}{x!}\int_{\mu}^{\infty}e^{-z}z^x\,dz$$

で与えられることを示せ．

6.21 Neyman (1939) 型伝播性分布　x をポアソン分布 $Po(\mu_1)$ に従う確率変数，y を，その条件つき確率変数 $y|x$ がポアソン分布 $Po(\mu_2 x)$ に従う確率変数とする．y の特性関数は

$$\varphi(t) = \exp\{-\mu_1[1 - e^{-\mu_2(1-e^{it})}]\}$$

で

$$\mathscr{E}(y) = \mu_1\mu_2, \qquad \sigma^2(y) = \mu_1\mu_2(1+\mu_2)$$

となることを示せ．

6.22　x は 2 項分布 $Bi(n,p)$ を持つ確率変数とする．このとき，x の c.d.f. は

$$(n-x)\binom{n}{x}\int_0^q y^{n-x-1}(1-y)^x\,dy$$

で与えられることを示せ．

6.23 血液検査の問題　大母集団から k 人を一組にして，次のように血液検査を行なう．k 人から血液をとり，それを集める．集められた血液が陰性ならば，k 人についてただ 1 回だけの血液検査で十分である．陽性となれば，各人の血液検査が別々に行なわれる．ランダムに選ばれた 1 人の検査で陰性となる確率を q とする．検査の全回数の期待値を最小にする k の値は，区間

$$\left(\sqrt[k]{\frac{1}{k(k+1)(1-q)}},\ \sqrt[k]{\frac{q}{k(k-1)(1-q)}}\right)$$

が q を含むような正の整数であることを示せ．q が 1 に近いとき

$$k^2 g^k \log g + 1 = 0 \qquad \text{[Dorfman (1943)]}$$

の解が k の近似値になることを示せ．

6.24　$1,2,\cdots,n$ と番号をつけた札が壺の中にはいっている．札を取り出してこれをもとに戻す．この操作を，1 度取り出した札が再度出るまで行なう．x を，取り出した札がふたたび取り出されるまでの回数を表わす確率変数とする．x の p.f. は

$$p(x) = (x-1)!\binom{n}{x-1}\frac{(x-1)}{n^k}, \qquad x = 2,\cdots,n+1$$

で与えられることを示せ．また

$$\sum_{x=2}^{n+1} p(x) = 1$$

を確かめよ．

$$\mathscr{E}(x) = 2 + \left(1-\frac{1}{n}\right) + \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \cdots + \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-1}{n}\right)$$

を示せ．

6.25　6.5 節の 2 項待ち時間問題において，k 個続けて C を得るまでの試行の回数を x で表わす．x の m.g.f. は

$$\phi(t) = (pe^t)^k(1-pe^t)(1-e^t+p^kqe^{(k+1)t})^{-1}$$

で

$$\mathscr{E}(x) = \frac{1-p^k}{p^kq}$$

となることを示せ．

6.26 クラスの大きさの分布に関する問題 —— 多項分布の場合 ——　(6.3.3) の p.f. を持つ多項分布において

$$p_1 = \cdots = p_{k+1} = \frac{1}{k+1}$$

とする．$r_0, r_1, \cdots, r_n$ を $(x_1,\cdots,x_{k+1})$ の成分の中で，それぞれ $0,1,\cdots,n$ となる回数とする．$(r_0,r_1,\cdots,r_n)$ の p.f. は

$$\frac{n!(k+1)!(k+1)^{-n}}{(0!)^{r_0}(1!)^{r_1}\cdots(n!)^{r_n}r_0!\,r_1!\cdots r_n!}$$

で与えられ

$$r_0 + r_1 + \cdots + r_n = k+1, \qquad r_1 + 2r_2 + \cdots + nr_n = n$$

を満たすことを示せ．さらに

$$\mathscr{E}[r_0^{[s_0]}r_1^{[s_1]}\cdots r_n^{[s_n]}] = \frac{n!(k+1)!(k+1)^{-n}(k+1-B)^{n-A}}{(0!)^{s_0}(1!)^{s_1}\cdots(n!)^{s_n}(n-A)!(k+1-B)!}$$

を示せ．ただし，$A = s_1 + 2s_2 + \cdots + ns_n$，$B = s_0 + s_1 + \cdots + s_n$．これから $\mathscr{E}(r_t)$，$\sigma^2(r_t)$，$\mathrm{cov}(r_t, r_u)$ を求めよ．[Tukey (1949 *b*)]．

6.27 クラスの大きさの分布に関する問題 —— 超幾何分布の場合 ——　おのおの M 枚のカードからなる $(k+1)$ 種類のカードがある．$M(k+1)$ 枚のカードをよく混ぜ合わせて，n 枚を取り出す．r_0 を 0 カード（空白）の個数，r_1 を 1 カードの個数，r_2 を 2 カードの個数，…，とする．このとき $(n+1)$ 次元確率変数 $(r_0,r_1,\cdots,r_n)$ の p.f. は

$$\frac{(M!)^{k+1}(k+1)!\binom{M(k+1)}{n}^{-1}}{r_0!r_1!\cdots r_n![0!(M-0)!]^{r_0}[1!(M-1)!]^{r_1}\cdots[n!(M-n)!]^{r_n}}$$

で与えられることを示せ．さらに

$$\mathscr{E}[r_0^{[s_0]}r_1^{[s_1]}\cdots r_n^{[s_n]}] = \frac{(M!)^B(k+1)!\binom{M(k+1-B)}{n-A}}{(k+1-B)!\binom{M(k+1)}{n}[0!(M-0)!]^{s_0}\cdots[n!(M-n)!]^{s_n}}$$

を示せ．ただし，$A = s_1 + 2s_2 + \cdots + ns_n$，$B = s_0 + s_1 + \cdots + s_n$．これから $\mathscr{E}(r_t)$，$\sigma^2(r_t)$，$\mathrm{cov}(r_t, r_u)$ を求めよ．[Tukey (1949 *b*)]．

6.28 2 組のカードの組合せ問題　A を N 枚のカードからなる 1 組のカードとし，そのおのおのは $S_1,\cdots,S_k$ のどれかの組に属し，その枚数をそれぞれ $m_1,\cdots,m_k$ とする．B もやはり N 枚からなる 1 組のカードとし，$S_1,\cdots,S_k$ の組には，それぞれ $n_1,$

$\cdots, n_k$ 枚のカードがあるとする．A のカードをよく混ぜて，1 直線に表を出して並べる．B のカードも同様にして，A のカードの下に並べる．この置き換えで，上下の 1 対が同じ組にはいっている数を x で表わす．可能な置き換えに等確率を与えるとすれば，x の p.f. は

$$\Phi = \frac{1}{M}\left(\sum_{i,j=1}^{k} a_i b_j e^{\delta_{ij}t}\right)^N$$

を展開したときの $e^{tx}a_1^{m_1}\cdots a_k^{m_k}b_1^{n_1}\cdots b_k^{n_k}$ の係数に等しいことを示せ．ただし，$\delta_{ij}=0$ $(i \neq j)$，1 $(i=j)$ で

$$M = \frac{(N!)^2}{m_1!\cdots m_k!\,n_1!\cdots n_k!}.$$

x の r 次のモーメント μ_r' は

$$\left[\frac{\partial^r \Phi}{\partial t^r}\right]_{t=0}$$

を展開したときの，$a_1^{m_1}\cdots a_k^{m_k}b_1^{n_1}\cdots b_k^{n_k}$ の係数で与えられることを示せ．特に

$$\mathscr{E}(x) = \sum_{i=1}^{k}\frac{m_i n_i}{N}$$

$$\sigma^2(x) = \frac{1}{N^2(N-1)}\left[\left(\sum_{i=1}^{k} m_i n_i\right)^2 - N\sum_{i=1}^{k}(m_i^2 n_i + m_i n_i^2) + N^2\sum_{i=1}^{k} m_i n_i\right]$$

を示せ．[Battin (1942), Kaplansky と Riordan (1945)]．

第7章　連続型分布

本章では数理統計学で重要な連続型確率分布とその性質を示す．ここでの結果は後章で役立つであろう．

7.1　矩形分布

最も簡単な連続型分布は

(7.1.1)
$$f(x) = \begin{cases} \dfrac{1}{\omega}, & \mu - \dfrac{\omega}{2} \leqslant x \leqslant \mu + \dfrac{\omega}{2} \\ 0, & x < \mu - \dfrac{\omega}{2},\quad x > \mu + \dfrac{\omega}{2} \end{cases}$$

で定義される p.d.f. を持つ分布である．この p.d.f. を持つ確率分布を**矩形分布***) $R(\mu,\omega)$ と呼ぶ．この分布の平均と分散は

(7.1.2)
$$\mathscr{E}(x) = \mu, \qquad \sigma^2(x) = \frac{\omega^2}{12}.$$

パラメータ ω を分布の**範囲**という．

x が矩形分布 $R(\mu,\omega)$ に従う確率変数のとき

$$y = \frac{x - \mu + \frac{1}{2}\omega}{\omega}$$

は矩形分布 $R\left(\frac{1}{2},1\right)$ に従うことに注意せよ．y は区間 $[0,1]$ では 0 でない値を持つ．

矩形分布 $R\left(\frac{1}{2},1\right)$ を持つ確率変数の重要な場合は，次のようになる．

7.1.1　x が連続な c.d.f. $F(x)$ を持つ確率変数のとき，確率変数 $y=F(x)$ は矩

*) 一様分布ということもある．(訳注)

形分布 $R\left(\frac{1}{2},1\right)$ に従う.

これは y の c.d.f. は

$$H(y)=P(F(x)\leqslant y)=\begin{cases}1, & y>1\\ y, & 0<y\leqslant 1\\ 0, & y\leqslant 0\end{cases}$$

となるから，矩形分布 $R\left(\frac{1}{2},1\right)$ に従うことがわかる.

7.2 正 規 分 布

連続型確率分布の最も重要な分布関数は，**正規**または**ガウス分布関数**である．その p.d.f. は $-\infty<x<\infty$ なる x に対して

(7.2.1)
$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}$$

と書かれる．μ, σ^2 はパラメータである．この 2 つのパラメータは，実際には，正規分布の平均と分散であることがわかる.

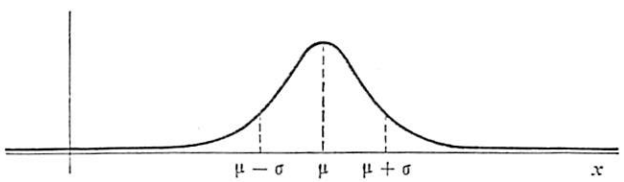

図 7.1　正規 p.d.f. (7.2.1) のグラフ.

(7.2.1) を p.d.f. に持つ分布を正規分布 $N(\mu,\sigma^2)$ という．あるいは，簡単に分布 $N(\mu,\sigma^2)$ と書く．(7.2.1) のグラフは図 7.1 のように，$x=\mu$ に関して対称で，$x=\mu$ で最大値 $1/\sqrt{2\pi}\sigma$ を持つ．$x=\mu\pm\sigma$ のときは変曲点を持つ.

正規分布を表にする場合，最も便利なのは，次の確率変数 y についてである．ただし，$y=(x-\mu)/\sigma$．y の p.d.f. を正規分布の標準形 $N(0,1)$ という.

$$P(x\leqslant x')=P(y\leqslant y')=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y'}e^{-\frac{1}{2}y^2}\,dy$$

になる．ただし，$y'=(x'-\mu)/\sigma$.

正規分布の標準形 c.d.f. $\Phi(x)$ は

(7.2.1 *a*)
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{1}{2}y^2}\,dy$$

で定義される．$\Phi(x)$ は表にまとめられているが，National Bureau of Standards (1942)，Greenwood と Hartley (1961) が特に広範囲にわたっている.

まず，正規分布関数の x 軸全体にわたる積分が 1 になることを示そう.

$y=(x-\mu)/\sigma$ と置くと

(7.2.2)
$$\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}\,dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}y^2}\,dy.$$

したがって，$1/\sqrt{2\pi}$ をのぞいた積分を I と置いて，$I=\sqrt{2\pi}$ を示せばよい．さて

(7.2.3)
$$I^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(y_1^2+y_2^2)}\,dy_1\,dy_2$$

において，

$$y_1=r\cos\theta,\qquad y_2=r\sin\theta$$

によって，極座標に変換すると

(7.2.3 *a*)
$$I^2=\int_0^{2\pi}\int_0^{\infty}re^{-\frac{1}{2}r^2}\,dr\,d\theta=2\pi$$

となる．ゆえに，$I=\sqrt{2\pi}$.

(7.2.1) の特性関数については

(7.2.4)
$$\varphi(t)=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^{itx-\frac{1}{2}(x-\mu)^2/\sigma^2}\,dx$$

より

(7.2.4 *a*)
$$\varphi(t)=e^{it\mu-\frac{1}{2}\sigma^2t^2}\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(x-\mu-i\sigma^2t)^2/\sigma^2}\,dx$$

となる．(7.2.4 *a*) の積分は $e^{-\frac{1}{2}z^2/\sigma^2}$（ただし，$z$ は複素変数）を複素平面上で実軸に平行に沿って積分したものである．すなわち，直線 $y=-i\sigma^2t$ にわたる積分である．これは，この関数を実軸に沿っての積分，すなわち $\int_{-\infty}^{\infty}e^{-\frac{1}{2}x^2/\sigma^2}\,dx$ に等しくなる．この値は $\sigma\sqrt{2\pi}$ である．よって

7.2.1　分布 $N(\mu,\sigma^2)$ の特性関数は

(7.2.5)
$$\varphi(t)=e^{it\mu-\frac{1}{2}\sigma^2t^2}.$$

微分によって，分布 $N(\mu,\sigma^2)$ の平均と分散が

(7.2.6) $$\mathscr{E}(x)=\mu,\qquad \sigma^2(x)=\sigma^2$$

として求められる．

特性関数 (7.2.5) において (μ,σ^2) をパラメータとみなすと，(5.3.7) を満たすから

7.2.2　分布 $N(\mu,\sigma^2)$ は (μ,σ^2) に関して再生的である．

事実，**7.2.2** より強い結果が次の **7.2.3** に述べられている．

7.2.3　x_1, x_2 を独立な確率変数で，それぞれ分布 $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$ を持つとき，c_1, c_2 をともには 0 にならない数として，$L=c_1x_1+c_2x_2$ とすると，L の分布は $N(c_1\mu_1+c_2\mu_2, c_1^2\sigma_1^2+c_2^2\sigma_2^2)$ になる．

これは L の特性関数を考えるとすぐわかる．(5.3.6) より

$$\varphi(t)=\varphi_1(c_1t)\cdot\varphi_2(c_2t).$$

ただし，$\varphi_1(t)$, $\varphi_2(t)$ はそれぞれ x_1, x_2 の特性関数である．

$$\varphi_j(c_jt)=e^{itc_j\mu_j-\frac{1}{2}c_j^2\sigma_j^2t^2},\qquad j=1,2$$

より

(7.2.7) $$\varphi(t)=\exp\left[i(c_1\mu_1+c_2\mu_2)t-\frac{1}{2}(c_1^2\sigma_1^2+c_2^2\sigma_2^2)t^2\right].$$

これは (7.2.5) により，正規分布 $N(c_1\mu_1+c_2\mu_2, c_1^2\sigma_1^2+c_2^2\sigma_2^2)$ の特性関数である．**5.1.3** の議論によって，**7.2.3** が成り立つことがわかる．この定理は，それぞれ正規分布に従う互いに独立な k 個の確率変数についても成り立つ．これは読者自身試みよ．

7.3　2 変数正規分布

正規分布は確率論および統計学ではきわめて重要な分布だから，k 変数に進む前に，2 変数について，ある程度詳しく議論する必要がある．2 変数正規分布の p.d.f. は

(7.3.1) $$f(x_1,x_2)=\frac{\sqrt{|\sigma^{ij}|}}{2\pi}e^{-\frac{1}{2}Q(x_1,x_2)}$$

と書くのが最も便利であろう．ただし，(x_1,x_2) は R_2 の任意の点で

(7.3.2) $$Q(x_1,x_2)=\sigma^{11}(x_1-\mu_1)^2+2\sigma^{12}(x_1-\mu_1)(x_2-\mu_2)+\sigma^{22}(x_2-\mu_2)^2.$$

μ_1, μ_2 はそれぞれ x_1, x_2 の平均で，行列 $\|\sigma^{ij}\|$, $i,j=1,2$ (ただし $\sigma^{12}=\sigma^{21}$) は，

(3.5.3) で定義された x_1, x_2 の共分散行列 $\|\sigma_{ij}\|$ の逆行列であることがあとでわかる．もちろん，$\|\sigma_{ij}\|$ は正定値であるとする．したがって $\|\sigma^{ij}\|$ も正定値，すなわち $Q(x_1,x_2)$ は正値 2 次形式である．

(7.3.1) を p.d.f. に持つ 2 変数正規分布を簡単に，$N(\{\mu_i\},\|\sigma_{ij}\|)$, $i,j=1,2$ と書く．

まず

(7.3.3) $$\int_{R_2}f(x_1,x_2)\,dx_1\,dx_2=1$$

すなわち

(7.3.4) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{1}{2}Q(x_1,x_2)}\,dx_1\,dx_2=\frac{2\pi}{\sqrt{|\sigma^{ij}|}}$$

を示そう．

$y_1=x_1-\mu_1$, $y_2=x_2-\mu_2$ と置くと，(7.3.4) の左辺は

(7.3.5) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}[\sqrt{\sigma^{11}}(y_1+\sigma^{12}/\sigma^{11}y_2)]^2-\frac{1}{2}[y_2\sqrt{|\sigma^{ij}|/\sigma^{11}}]^2\right\}dy_1\,dy_2$$

となる．変換

(7.3.6) $$z_1=\sqrt{\sigma^{11}}\left(y_1+\frac{\sigma^{12}}{\sigma^{11}}y_2\right),\qquad z_2=\sqrt{\frac{|\sigma^{ij}|}{\sigma^{11}}}\,y_2$$

のヤコビアンは

$$\left|\frac{\partial(y_1,y_2)}{\partial(z_1,z_2)}\right|=\left|\frac{\partial(z_1,z_2)}{\partial(y_1,y_2)}\right|^{-1}=\frac{1}{\sqrt{|\sigma^{ij}|}}$$

だから，(7.3.4) の左辺は

$$\frac{1}{\sqrt{|\sigma^{ij}|}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(z_1^2+z_2^2)}\,dz_1\,dz_2$$

になる．しかも (7.2.3), (7.2.3 a) より，この 2 重積分は 2π であるから，(7.3.4) が成立する．

μ_1, μ_2 が x_1, x_2 の平均であることを示すには，まず，(7.3.3) の両辺を μ_1, μ_2 で微分して

$$\mathscr{E}[\sigma^{11}(x_1-\mu_1)+\sigma^{12}(x_2-\mu_2)]=0$$
$$\mathscr{E}[\sigma^{21}(x_1-\mu_1)+\sigma^{22}(x_2-\mu_2)]=0.$$

ゆえに，

(7.3.7) $$\sigma^{11}\mathscr{E}(x_1-\mu_1)+\sigma^{12}\mathscr{E}(x_2-\mu_2)=0$$
$$\sigma^{21}\mathscr{E}(x_1-\mu_1)+\sigma^{22}\mathscr{E}(x_2-\mu_2)=0.$$

これは $\mathscr{E}(x_1-\mu_1)$, $\mathscr{E}(x_2-\mu_2)$ に関する斉次 1 次方程式である．よって，$|\sigma^{ij}|\neq 0$ より，一意解

(7.3.8)
$$\mathscr{E}(x_1-\mu_1)=0,\qquad \mathscr{E}(x_2-\mu_2)=0$$

を持つ．(7.3.8) は

(7.3.9)
$$\mu_1=\mathscr{E}(x_1),\qquad \mu_2=\mathscr{E}(x_2)$$

に他ならない．ゆえに，μ_1, μ_2 はそれぞれ x_1, x_2 の平均である．

分散，共分散についても考えよう．(7.3.4) の両辺を σ^{11} で微分して，両辺に $-\sqrt{|\sigma^{ij}|}/\pi$ を掛けると

$$\frac{\sqrt{|\sigma^{ij}|}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-\mu_1)^2 e^{-\frac{1}{2}Q(x_1,x_2)}\,dx_1\,dx_2=\frac{\sigma^{22}}{|\sigma^{ij}|}.$$

しかし，この左辺は x_1 の分散 $\mathscr{E}(x_1-\mu_1)^2$ である．この分散を σ_{11} とすると

$$\sigma_{11}=\frac{\sigma^{22}}{|\sigma^{ij}|}.$$

x_1 と x_2 の共分散を σ_{12}, x_2 の分散を σ_{22} とすれば，同様に

$$\sigma_{12}=-\frac{\sigma^{12}}{|\sigma^{ij}|},\qquad \sigma_{22}=\frac{\sigma^{11}}{|\sigma^{ij}|}.$$

よって，共分散行列 $\|\sigma_{ij}\|$ は (7.3.2) の σ^{11}, σ^{12}, σ^{22} で表わされ，

(7.3.10)
$$\|\sigma_{ij}\|=\|\sigma^{ij}\|^{-1}$$

となる．

まとめると

7.3.1　(7.3.2) の定数 μ_1, μ_2 は x_1, x_2 の平均で，行列 $\|\sigma^{ij}\|$ は x_1, x_2 の共分散行列 $\|\sigma_{ij}\|$ の逆行列である．

x_1, x_2 の分散を σ_1^2, σ_2^2, 共分散を $\sigma_1\sigma_2\rho$（ただし ρ は x_1 と x_2 との相関係数）で表わすと，2 変数正規分布の p.d.f. (7.3.1) は

(7.3.11)
$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\,e^{-\frac{1}{2}Q(x_1,x_2)}$$

とも書ける．ただし

(7.3.12)
$$Q(x_1,x_2)=\frac{1}{(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right].$$

正規分布の p.d.f. は楕円 $Q(x_1,x_2)=$ 定数 上では，一定な値をとり，$f(x_1,x_2)$ はこの分布の中心すなわち (μ_1,μ_2) で最大になる．

さて，(x_1,x_2) の特性関数すなわち

(7.3.13)
$$\varphi(t_1,t_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{it_1x_1+it_2x_2}f(x_1,x_2)\,dx_1\,dx_2$$

を決めよう．ふたたび，$y_1=x_1-\mu_1$, $y_2=x_2-\mu_2$ とすると

(7.3.14)
$$\varphi(t_1,t_2)=\frac{e^{i\mu_1t_1+i\mu_2t_2}}{2\pi\sqrt{|\sigma_{ij}|}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{1}{2}Q'(y_1,y_2)}\,dy_1\,dy_2.$$

ただし，$Q'(y_1,y_2)=\sigma^{11}y_1^2+\sigma^{22}y_2^2+2\sigma^{12}y_1y_2-2it_1y_1-2it_2y_2$．しかし，$Q'(y_1,y_2)$ は

(7.3.15)
$$Q'(y_1,y_2)=z_1^2+z_2^2+(\sigma_{11}t_1^2+2\sigma_{12}t_1t_2+\sigma_{22}t_2^2)$$

となる．ここで，y と z を関連づける変換

$$z_1=\sqrt{\sigma^{11}}\left(y_1+\frac{\sigma^{12}}{\sigma^{11}}y_2-\frac{it_1}{\sigma^{11}}\right),\qquad z_2=\sqrt{\frac{|\sigma^{ij}|}{\sigma^{11}}}\left(y_2+\frac{i\sigma^{12}t_1-i\sigma^{11}t_2}{|\sigma^{ij}|}\right)$$

のヤコビアンは

$$\left|\frac{\partial(y_1,y_2)}{\partial(z_1,z_2)}\right|=1/\sqrt{|\sigma^{ij}|}$$

である．(7.2.3), (7.2.3 a) を用いると，(7.3.14) より次が成り立つ．

7.3.2　2 変数正規分布 (7.3.1)［または (7.3.11)］の特性関数は

(7.3.16)
$$\varphi(t_1,t_2)=\exp\left[i(\mu_1t_1+\mu_2t_2)-\frac{1}{2}(\sigma_{11}t_1^2+\sigma_{22}t_2^2+2\sigma_{12}t_1t_2)\right]$$

である．

読者は，2 変数正規分布が平均と共分散行列とのベクトルに関して再生的であることを確かめよ．

z_1, z_2 がともに複素変数のとき，(7.3.14) において，y を z に置き換えて得られる積分 $\int e^{-\frac{1}{2}z_1^2}dz_1$, $\int e^{-\frac{1}{2}z_2^2}dz_2$ は複素平面上で実軸に平行な直線にそっての積分である．しかし，この積分は実軸にそっての積分に等しい．

特性関数における t_1, t_2 に関する 2 次形式の行列は，共分散行列 $\|\sigma_{ij}\|$ である．一方，p.d.f. (7.3.1) における x_1, x_2 に関する 2 次形式の行列は，共分散行列の逆行列すなわち $\|\sigma^{ij}\|$ である．

読者は (7.3.16) に微分演算を施して，x_1, x_2 の平均がそれぞれ μ_1, μ_2 で，その共分散行列が $\|\sigma_{ij}\|$ であることを確かめよ．

(7.3.16) で $t_2=0$ と置くと，x_1 の特性関数，すなわち

(7.3.17)
$$\varphi(t_1,0)=e^{i\mu_1t_1-\frac{1}{2}\sigma_{11}t_1^2}$$

を得る．したがって，**7.2.1** から次がいえる．

7.3.3　分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, 2$ において，x_1 の周辺分布は $N(\mu_1, \sigma_{11})$ である．

もちろん，x_2 の周辺分布についても同じ事柄が成り立つ．

一般的には，c_1, c_2 をともには 0 でない実数とすると，$L = c_1x_1 + c_2x_2$ の特性関数は

$$\varphi(c_1t, c_2t) = \exp\left[i(c_1\mu_1 + c_2\mu_2)t - \frac{1}{2}(\sigma_{11}c_1^2 + 2\sigma_{12}c_1c_2 + \sigma_{22}c_2^2)t^2\right] \tag{7.3.18}$$

で与えられる．よって，次を得る．

7.3.4　(x_1, x_2) が分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, 2$ に従うとき，$L = c_1x_1 + c_2x_2$ の分布は次式で与えられる．

$$N\left(\sum_{i=1}^{2} c_i\mu_i,\ \sum_{i,j=1}^{2} \sigma_{ij}c_ic_j\right).$$

7.2.3 は **7.3.4** において，共分散 $\sigma_{12} = 0$, $\sigma_{11} = \sigma_1^2$, $\sigma_{22} = \sigma_2^2$ とした場合に過ぎない．

(x_1, x_2) が 2 変数正規分布を持つとき，条件つき確率変数 $x_2|x_1$ について考えよう．

x_1 の周辺 p.d.f. は

$$f_1(x_1) = \frac{1}{\sqrt{2\pi}\,\sigma_1} e^{-\frac{1}{2}(x_1-\mu_1)^2/\sigma_1^2}. \tag{7.3.19}$$

ただし，$\sigma_1^2 = \sigma_{11} = \sigma^2(x_1)$. (2.9.13) より，(7.3.11) の 2 変数正規分布の p.d.f. を用いると

$$f(x_2|x_1) = \frac{1}{\sqrt{2\pi}\,\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x_1,x_2)+\frac{1}{2}(x_1-\mu_1)^2/\sigma_1^2}.$$

ただし，$Q(x_1, x_2)$ は (7.3.12) で与えられている．結局，

$$f(x_2|x_1) = \frac{1}{\sqrt{2\pi}\,\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left[(x_2-\mu_2) - \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\right]^2\right\} \tag{7.3.20}$$

となるから，**7.3.5** が成り立つ．

7.3.5　(7.3.11) を p.d.f. に持つ 2 変数正規分布において，条件つき確率変数 $x_2|x_1$ は分布

$$N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1), \sigma_2^2(1-\rho^2)\right)$$

に従う．

対称性によって，$x_1|x_2$ の条件つき分布についても同様なことが成り立つ．

$x_2|x_1$ の平均，すなわち (7.3.20) における x_2 の x_1 上への回帰関数は x_1 の線形関数

$$\mu(x_2|x_1) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1).$$

よって，x_2 の x_1 への回帰直線の方程式は

$$x_2 = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1) \tag{7.3.21}$$

である．これは (3.8.6) による x_2 の x_1 への最小 2 乗回帰直線と同じである．この関係は x_1 の x_2 への回帰関数にも成り立つ．したがって，3.8 (a) 節における注意は次のように明確に述べられる．

7.3.6　2 変数正規分布において，2 つの回帰関数はともに線形で，最小 2 乗法による回帰直線に一致する．さらに，条件つき確率変数 $x_2|x_1$, $x_1|x_2$ の分散は，それぞれ (3.8.8) で定義された最小 2 乗残差分散 $\sigma_{2\cdot1}^2$, $\sigma_{1\cdot2}^2$ に等しい．

7.4　k 変数正規分布

(a)　k 変数正規分布の p.d.f.

多変数または k 変数正規分布およびその性質は 7.3 節の $k=2$ の場合を拡張して得られる．

k 変数正規分布の p.d.f. は

$$f(x_1, \cdots, x_k) = \frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}} e^{-\frac{1}{2}Q(x_1,\cdots,x_k)} \tag{7.4.1}$$

である．ここで $(x_1, \cdots, x_k)$ は R_k の点で，$Q(x_1, \cdots, x_k)$ は

$$Q(x_1, \cdots, x_k) = \sum_{i,j=1}^{k} \sigma^{ij}(x_i-\mu_i)(x_j-\mu_j). \tag{7.4.2}$$

$\|\sigma^{ij}\|$ は共分散行列 $\|\sigma_{ij}\|$ の逆行列，μ_i は x_i の平均である．

$\|\sigma^{ij}\|$ は正定値を仮定しておくと，$|\sigma^{ij}| \neq 0$ となり，$|\sigma_{ij}| \neq 0$ である．(7.4.1) を p.d.f. として持つ分布を **k 変数正規分布**といい，$N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, \cdots, k$ と書く．確率変数 $(x_1, \cdots, x_k)$ の標本空間は k 次元ユークリッド空間 R_k である．

$y_i = x_i - \mu_i$, $i = 1, \cdots, k$ とする．$(x_1, \cdots, x_k)$ の R_k にわたる積分が 1 であること

を示すには

$$(7.4.3) \qquad \int_{R_k} \exp\left(-\frac{1}{2}\sum_{i,j=1}^{k} \sigma^{ij} y_i y_j\right) dy_1 \cdots dy_k = \frac{(2\pi)^{\frac{1}{2}k}}{\sqrt{|\sigma^{ij}|}}$$

を示せばよい.

$$\sum_{i,j=1}^{k} \sigma^{ij} y_i y_j = \sigma^{11}\left(y_1 + \frac{\sum_{j=2}^{k} \sigma^{1j} y_j}{\sigma^{11}}\right)^2 + \sum_{i,j=2}^{k}\left(\sigma^{ij} - \frac{\sigma^{1i}\sigma^{1j}}{\sigma^{11}}\right) y_i y_j$$

より,

$$z_1 = \sqrt{\sigma^{11}}\left(y_1 + \frac{1}{\sigma^{11}}\sum_{j=2}^{k} \sigma^{1j} y_j\right), \qquad \sigma^{ij}_{(1)} = \sigma^{ij} - \frac{\sigma^{1i}\sigma^{1j}}{\sigma^{11}}, \qquad i, j = 2, \cdots, k$$

とすると

$$\sum_{i,j=1}^{k} \sigma^{ij} y_i y_j = z_1^2 + \sum_{i,j=2}^{k} \sigma^{ij}_{(1)} y_i y_j.$$

この方法を続ける. すなわち

$$\sigma^{ij}_{(p)} = \sigma^{ij}_{(p-1)} - \frac{\sigma^{pi}_{(p-1)}\sigma^{pj}_{(p-1)}}{\sigma^{pp}_{(p-1)}}, \qquad i, j = p+1, \cdots, k\ ;\ p = 1, \cdots, k-1$$

$$\sigma^{ij}_{(0)} = \sigma^{ij}, \qquad i, j = 1, \cdots, k$$

を用いると

$$\sum_{i,j=1}^{k} \sigma^{ij} y_i y_j = \sum_{i=1}^{k} z_i^2$$

になる. ここで

$$(7.4.4) \qquad z_i = \sqrt{\sigma^{ii}_{(i-1)}}\left(y_i + \frac{1}{\sigma^{ii}_{(i-1)}}\sum_{j=i+1}^{k} \sigma^{ij}_{(i-1)} y_j\right), \qquad i = 1, \cdots, k.$$

$\|\sigma^{ij}\|$ が正定値であるから, $\sigma^{11}, \sigma^{22}_{(1)}, \cdots, \sigma^{kk}_{(k-1)}$ はすべて正である.

この方法は (7.4.3) の指数の正定値 2 次形式を 2 乗和に線形変換したもので, ラグランジュの方法としてよく知られている. もちろん, このように 2 乗和に変換するためのいろいろな線形変換が考えられるだろう.

変換 (7.4.4) のヤコビアンは

$$\left|\frac{\partial(y_1, \cdots, y_k)}{\partial(z_1, \cdots, z_k)}\right| = 1 \Big/ \left|\frac{\partial(z_1, \cdots, z_k)}{\partial(y_1, \cdots, y_k)}\right| = \frac{1}{\sqrt{\sigma^{11}\sigma^{22}_{(1)}\cdots\sigma^{kk}_{(k-1)}}}$$

であるから, (7.4.3) の左辺は

$$(7.4.5) \qquad \frac{1}{\sqrt{\sigma^{11}\sigma^{22}_{(1)}\cdots\sigma^{kk}_{(k-1)}}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} z_i^2\right) dz_1 \cdots dz_k$$

になる. (7.2.3 a) を用いると, (7.4.5) の k 重積分は値 $(2\pi)^{-\frac{1}{2}k}$ を持つ. したがって,

$$(7.4.6) \qquad \sigma^{11}\sigma^{22}_{(1)}\cdots\sigma^{kk}_{(k-1)} = |\sigma^{ij}|$$

を示すことになる. そのために, 行列式 $|\sigma^{ij}|$ を次のように評価する.

$$(7.4.7) \qquad |\sigma^{ij}| = \begin{vmatrix} \sigma^{11} & \sigma^{12} & \cdots & \sigma^{1k} \\ \sigma^{21} & \sigma^{22} & \cdots & \sigma^{2k} \\ \vdots & \vdots & & \vdots \\ \sigma^{k1} & \sigma^{k2} & \cdots & \sigma^{kk} \end{vmatrix} = \sigma^{11} \begin{vmatrix} 1 & \sigma^{12} & \cdots & \sigma^{1k} \\ \frac{\sigma^{21}}{\sigma^{11}} & \sigma^{22} & \cdots & \sigma^{2k} \\ \vdots & \vdots & & \vdots \\ \frac{\sigma^{k1}}{\sigma^{11}} & \sigma^{k2} & \cdots & \sigma^{kk} \end{vmatrix}.$$

右辺の行列の第 1 列に σ^{12} を掛けて, 第 2 列から引く. 次に, 第 1 列に σ^{13} を掛けて, 第 3 列から引く. これを繰り返すと

$$(7.4.8) \qquad |\sigma^{ij}| = \sigma^{11} \begin{vmatrix} \sigma^{22}_{(1)} & \sigma^{23}_{(1)} & \cdots & \sigma^{2k}_{(1)} \\ \sigma^{32}_{(1)} & \sigma^{33}_{(1)} & \cdots & \sigma^{3k}_{(1)} \\ \vdots & \vdots & & \vdots \\ \sigma^{k2}_{(1)} & \sigma^{k3}_{(1)} & \cdots & \sigma^{kk}_{(1)} \end{vmatrix}.$$

これを $(k-1)$ 回続けて行なうと (7.4.6), すなわち

$$|\sigma^{ij}| = \sigma^{11}\sigma^{22}_{(1)}\cdots\sigma^{kk}_{(k-1)}$$

が示される. μ_i が p.d.f. (7.4.1) を持つ確率変数の平均であることを示すには, 2 次元の場合と同様に

$$\frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}}\int_{R_k} e^{-\frac{1}{2}Q(x_1, \cdots, x_k)}\, dx_1 \cdots dx_k = 1$$

を μ_i, $i = 1, \cdots, k$ で微分すると

$$\mathscr{E}\left[\sum_{j=1}^{k} \sigma^{ij}(x_j - \mu_j)\right] = 0, \qquad i = 1, \cdots, k$$

が得られる. ゆえに

$$(7.4.9) \qquad \sum_{j=1}^{k} \sigma^{ij}\mathscr{E}(x_j - \mu_j) = 0.$$

$\|\sigma_{ij}\|$ は正定値であると仮定しているので, $\|\sigma^{ij}\|$ も正定値であるから $|\sigma^{ij}| \neq 0$. ゆえに, (7.4.9) の唯一の解は

$$(7.4.10) \qquad \mathscr{E}(x_i - \mu_i) = 0, \quad \text{すなわち,} \quad \mathscr{E}(x_i) = \mu_i, \qquad i = 1, \cdots, k$$

を満たす. よって, (7.4.1) の $\mu_1, \cdots, \mu_k$ はそれぞれ $x_1, \cdots, x_k$ の平均である.

$\|\sigma^{ij}\|$ が共分散行列の逆行列であることを確かめるには，次のようにする．

$$\int_{R_k} e^{-\frac{1}{2}Q(x_1,\cdots,x_k)}dx_1\cdots dx_k = \frac{(2\pi)^{\frac{1}{2}k}}{\sqrt{|\sigma^{ij}|}} \tag{7.4.11}$$

の両辺を σ^{ij} で微分して，$-(1+\delta_{ij})\dfrac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}}$ を両辺に掛けると

$$\mathcal{E}[(x_i-\mu_i)(x_j-\mu_j)] = \sigma_{ij} \tag{7.4.12}$$

を得る．ただし δ_{ij} はクロネッカーの記号 δ. (7.4.12) は $i \fallingdotseq j$ だけでなく，$i=j$ に対しても成り立つ．したがって，次の結果が成り立つ．

7.4.1　(7.4.1) に含まれている定数 $\mu_1,\cdots,\mu_k$ は $x_1,\cdots,x_k$ の平均で，$\|\sigma^{ij}\|$ は共分散行列 $\|\sigma_{ij}\|$ の逆行列である．

x_i の分散を σ_i^2，x_i と x_j の共分散を $\sigma_i\sigma_j\rho_{ij}$ で表わすと，(7.4.1) は，少々煩雑だが (7.3.12) の型に書き直せる．これは読者自身で試みよ．

(b)　k 変数正規分布の特性関数

k 変数正規分布の特性関数は

$$\varphi(t_1,\cdots,t_k) = \exp\left\{\mathrm{i}\sum_{i=1}^k \mu_i t_i\right\}\cdot\frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}} \tag{7.4.13}$$

$$\cdot\int_{R_k}\exp\left\{-\frac{1}{2}\left[Q(x_1,\cdots,x_k)-2\mathrm{i}\sum_{i=1}^k(x_i-\mu_i)t_i\right]\right\}dx_1\cdots dx_k$$

になる．ただし，$Q(x_1,\cdots,x_k)$ は (7.4.2) で与えられている．

(7.4.13) を評価するのに，(7.2.4) から (7.2.4 a) を導いたように，[　] 内を変形する．まず，$y_i = x_i-\mu_i$，$i=1,\cdots,k$ と置くと

$$\sum_{i,j=1}^k \sigma^{ij}\left[y_i-\mathrm{i}\sum_{h=1}^k\sigma_{ih}t_h\right]\left[y_j-\mathrm{i}\sum_{g=1}^k\sigma_{gj}t_g\right] \tag{7.4.14}$$

$$\equiv \sum_{i,j=1}^k\sigma^{ij}y_iy_j - \mathrm{i}\sum_{i,j,h=1}^k\sigma^{ij}\sigma_{ih}t_hy_j - \mathrm{i}\sum_{i,j,g=1}^k\sigma^{ij}\sigma_{gj}t_gy_i$$

$$-\sum_{i,j,g,h=1}^k\sigma^{ij}\sigma_{ih}\sigma_{gj}t_gt_h.$$

$\sigma^{ij}=\sigma^{ji}$，もちろん，$\sigma_{ij}=\sigma_{ji}$ より $\sum_{i=1}^k\sigma^{ij}\sigma_{ih}$，$\sum_{j=1}^k\sigma^{ij}\sigma_{gj}$ はそれぞれクロネッカー記号 δ_{jh}，δ_{ig} に等しい．したがって，(7.4.14) の第 2 項は

$$Q(x_1,\cdots,x_k)-2\mathrm{i}\sum_{i=1}^k t_iy_i - \sum_{i,j=1}^k\sigma_{ij}t_it_j$$

となる．ゆえに

$$Q(x_1,\cdots,x_k)-2\mathrm{i}\sum_{i=1}^k t_iy_i = Q'(x_1,\cdots,x_k)+\sum_{i,j=1}^k\sigma_{ij}t_it_j. \tag{7.4.15}$$

ただし，$Q'(x_1,\cdots,x_k)$ は (7.4.14) での最初の項の 2 次形式である．よって，(7.4.13) は

$$\varphi(t_1,\cdots,t_k) = \exp\left(\mathrm{i}\sum_{i=1}^k\mu_it_i-\frac{1}{2}\sum_{i,j=1}^k\sigma_{ij}t_it_j\right)\cdot H \tag{7.4.16}$$

と書ける．ただし

$$H = \frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}}\int_{R_k}e^{-\frac{1}{2}Q'(x_1,\cdots,x_k)}dx_1\cdots dx_k.$$

$Q'(x_1,\cdots,x_k)$ は複素変数 $(x_i-\mu_i)+\mathrm{i}B_i$ $(i=1,\cdots,k)$ の $\|\sigma^{ij}\|$ に関する 2 次形式である．ここで B_i は実数である．$Q(x_1,\cdots,x_k)$ は $Q'(x_1,\cdots,x_k)$ において $B_i=0$ とした場合である．$x_i-\mu_i+\mathrm{i}B_i$ を y_i' に置き換えると，y_i' を y とみなせば，変換 (7.4.4) より，$Q'(x_1,\cdots,x_k)$ は 2 乗和 $\sum_{i=1}^k z_i^2$ になる．もちろん，z_i は複素変数で H の中の積分は，各 z_i のその複素平面上での実軸に平行な直線上の積分である．しかし，1 変数および 2 変数正規分布の場合に示したように，この積分は実軸上の積分に等しい．したがって，H は $Q'(x_1,\cdots,x_k)$ を $Q(x_1,\cdots,x_k)$ に置き換えたときの積分値，すなわち 1 に等しい．よって

7.4.2　分布 $N(\{\mu_i\},\|\sigma_{ij}\|)$，$i,j=1,\cdots,k$ の特性関数は

$$\varphi(t_1,\cdots,t_k) = \exp\left(\mathrm{i}\sum_{i=1}^k\mu_it_i-\frac{1}{2}\sum_{i,j=1}^k\sigma_{ij}t_it_j\right) \tag{7.4.17}$$

である．

特に，$t_{k_1+1}=\cdots=t_k=0$ とすると，特性関数は

$$\varphi(t_1,\cdots,t_{k_1},0,\cdots,0) = \exp\left(\mathrm{i}\sum_{i=1}^{k_1}\mu_it_i-\frac{1}{2}\sum_{i,j=1}^{k_1}\sigma_{ij}t_it_j\right) \tag{7.4.18}$$

となるから

7.4.3　$(x_1,\cdots,x_k)$ は k 変数正規分布 $N(\{\mu_i\},\|\sigma_{ij}\|)$，$i,j=1,\cdots,k$ を持つ確率変数とする．このとき，$(x_1,\cdots,x_{k_1})$ $(k_1<k)$ の周辺分布は k_1 変数正規分布

$N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, \cdots, k_1$ になる．

(c)　正規変数の線形関数の分布

(7.4.17) で $t_i = c_i t$, $i = 1, \cdots, k$ (ただし，c_i はすべては 0 でない) とすると，(5.3.6) で述べたように，線形関数 $L = \sum_{i=1}^{k} c_i x_i$ の c.f. $\varphi(t)$ は

(7.4.19)
$$\varphi(t) = \exp\left[\mathrm{i}\left(\sum_{i=1}^{k} c_i \mu_i\right)t - \frac{1}{2}\left(\sum_{i=1}^{k} \sigma_{ij} c_i c_j\right)t^2\right]$$

になる．**7.2.1** を用いると

7.4.4　$(x_1, \cdots, x_k)$ が k 変数分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, \cdots, k$ に従うならば，$L = c_1 x_1 + \cdots + c_k x_k$ は次の分布を持つ．

$$N\left(\sum_{i=1}^{k} c_i \mu_i, \sum_{i,j=1}^{k} \sigma_{ij} c_i c_j\right).$$

同様に，

7.4.5　一般に，$L_p = c_{p1} x_1 + \cdots + c_{pk} x_k$, $p = 1, \cdots, s$, $s \leqslant k$ を 1 次独立な確率変数とする．ただし，$(x_1, \cdots, x_k)$ は分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, \cdots, k$ に従う．このとき，$(L_1, \cdots, L_s)$ は次の s 次元分布を持つ．

$$N\left(\left\{\sum_{i=1}^{k} c_{pi} \mu_i\right\}, \left\|\sum_{i,j=1}^{k} \sigma_{ij} c_{pi} c_{qj}\right\|\right), \qquad p, q = 1, \cdots, s.$$

(d)　k 変数正規分布の条件つき分布

最後に k 変数正規分布における条件つき確率変数 $x_k | x_1, \cdots, x_{k-1}$ の p.d.f. を考えよう．定義より

(7.4.20)
$$f(x_k | x_1, \cdots, x_{k-1}) = \frac{f(x_1, \cdots, x_k)}{f_{12\cdots(k-1)}(x_1, \cdots, x_{k-1})}.$$

ここで，$f(x_1, \cdots, x_k)$ は (7.4.1) の p.d.f. で，$f_{12\cdots(k-1)}(x_1, \cdots, x_{k-1})$ は $(x_1, \cdots, x_{k-1})$ の周辺分布の p.d.f. である．すなわち，

(7.4.21)
$$f_{12\cdots(k-1)}(x_1, \cdots, x_{k-1}) = \frac{\sqrt{|\sigma_{(k)}^{pq}|}}{(2\pi)^{\frac{1}{2}(k-1)}} \exp\left[-\frac{1}{2}\sum_{p,q=1}^{k-1} \sigma_{(k)}^{pq}(x_p - \mu_p)(x_q - \mu_q)\right].$$

$\|\sigma_{(k)}^{pq}\|$ は $\|\sigma_{pq}\|$, $p, q = 1, \cdots, k-1$ の逆行列である．$\|\sigma_{ij}\|$ およびこの逆行列 $\|\sigma^{ij}\|$ については $i, j = 1, \cdots, k$ であることに注意せよ．

さて，$y_i = x_i - \mu_i$, $i = 1, \cdots, k$ とすると

(7.4.22)
$$\sum_{i,j=1}^{k} \sigma^{ij} y_i y_j \equiv \sigma^{kk}\left[y_k - \sum_{p=1}^{k-1} \beta_p^* y_p\right]^2 + \sum_{p,q=1}^{k-1} \sigma_{(k)}^{pq} y_p y_q$$

が確かめられる．ただし

(7.4.23)
$$\beta_p^* = \sum_{q=1}^{k-1} \sigma_{(k)}^{pq} \sigma_{qk}, \qquad p = 1, \cdots, k-1.$$

さらに，

(7.4.24)
$$\frac{|\sigma^{ij}|}{|\sigma_{(k)}^{pq}|} = \frac{|\sigma_{pq}|}{|\sigma_{ij}|} = \sigma^{kk}.$$

ただし，$i, j = 1, \cdots, k$, $p, q = 1, \cdots, k-1$.

(7.4.20) に $f(x_1, \cdots, x_k)$, $f_{1\cdots(k-1)}(x_1, \cdots, x_{k-1})$ を代入して，(7.4.22), (7.4.24) を用いると

(7.4.25)
$$f(x_k | x_1, \cdots, x_{k-1}) = \frac{\sqrt{\sigma^{kk}}}{\sqrt{2\pi}} \exp\left\{-\frac{\sigma^{kk}}{2}\left[(x_k - \mu_k) - \sum_{p=1}^{k-1} \beta_p^*(x_p - \mu_p)\right]^2\right\}.$$

したがって，次を得る．

7.4.6　$(x_1, \cdots, x_k)$ を，k 変数分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, \cdots, k$ を持つ確率変数とすれば，条件つき確率変数 $x_k | x_1, \cdots, x_{k-1}$ の分布は

(7.4.26)
$$N\left(\mu_k + \sum_{p=1}^{k-1} \beta_p^*(x_p - \mu_p), \frac{1}{\sigma^{kk}}\right)$$

である．ただし β_p^* は (7.4.23) で与えられている．

(7.4.25) から得られた平均 $\mu(x_k | x_1, \cdots, x_{k-1})$ は $x_1, \cdots, x_{k-1}$ に関して線形で，(3.8.18) で与えられた，x_k の $x_1, \cdots, x_{k-1}$ 上への最小 2 乗回帰平面に等しい．さらに，(7.4.25) における分散 $\sigma^2(x_k | x_1, \cdots, x_{k-1})$ は $1/\sigma^{kk}$. この値は，(3.8.25) で与えられたように，x_k の $x_1, \cdots, x_{k-1}$ 上への最小 2 乗残差分散に等しい．よって，次の結果が得られる．

7.4.7　k 変数正規分布においては，条件つき確率変数 $x_k | x_1, \cdots, x_{k-1}$ の平均と分散は，それぞれ x_k の $x_1, \cdots, x_{k-1}$ 上への最小 2 乗回帰関数およびその残差分散に等しい．

7.5 ガンマ分布

一般に，ガンマ関数 $\Gamma(g)$ は実数部分が正の複素数 g に対して，定積分

(7.5.1) $$\Gamma(g)=\int_0^\infty x^{g-1}e^{-x}\,dx$$

で定義される．しかし，実際は g が正の実数のとき有用である．このとき，部分積分によって

$$\Gamma(g)=(g-1)\Gamma(g-1)$$

となるから，正の整数 g に対しては

$$\Gamma(g)=(g-1)\,!.$$

$g>0$ で，g が整数でないときは

$$\Gamma(g)=(g-1)(g-2)\cdots\delta\Gamma(\delta).$$

ただし，$0<\delta<1$．特に，$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$．これは (7.2.3) の I を用いると，$I=\sqrt{2}\,\Gamma\left(\frac{1}{2}\right)$ となることから確かめられる．

後章で見るように，統計理論においては，g が $\frac{1}{2}$ の（正の）倍数のときが重要である．

ガンマ分布は正規確率変数の 2 乗和およびその 2 次形式を考察するときよく現われる．これは p. d. f.

(7.5.2) $$f(x)=\begin{cases}\dfrac{x^{\mu-1}e^{-x}}{\Gamma(\mu)}, & x>0\quad(\mu>0)\\ 0, & \text{その他}\end{cases}$$

を持つ分布である．この p. d. f. を持つ分布をガンマ分布 $G(\mu)$ と書く．Pearson (1906) の第 3 型分布ともいわれている．

(7.5.2) の r 次モーメント μ'_r は

(7.5.3) $$\mu'_r=\frac{1}{\Gamma(\mu)}\int_0^\infty x^{\mu+r-1}e^{-x}\,dx=\frac{\Gamma(\mu+r)}{\Gamma(\mu)}.$$

よって，

(7.5.4) $$\mathscr{E}(x)=\mu,\qquad \sigma^2(x)=\mu.$$

すなわち，ガンマ分布の平均と分散はともに μ に等しい．この性質はポアソン分布 (6.4.1) にもある．しかし，ポアソン分布は離散型分布である．

(7.5.2) で $\mu=p+1$ と置いたとき，関数

$$I(u,p)=\frac{1}{\Gamma(p+1)}\int_0^{u\sqrt{p+1}}x^pe^{-x}\,dx$$

を不完全ガンマ関数という．

Karl Pearson (1922) は $I(u,p)$ に対して u を 0 から 12，p を 0 から 50 まで動かしたときの表を作成している．

ガンマ分布 $G(\mu)$ の特性関数は

(7.5.5) $$\varphi(t)=\frac{1}{\Gamma(\mu)}\int_0^\infty x^{\mu-1}e^{-x(1-it)}\,dx$$

において，$x(1-it)=y$ とすると

(7.5.6) $$\varphi(t)=(1-it)^{-\mu}$$

になる．μ をガンマ分布のパラメータとみなせば，特性関数 (5.3.7) を満たすから

7.5.1 ガンマ分布 $G(\mu)$ は μ に関して再生的である．

すでに述べたように，ガンマ分布は正規確率変数の 2 次形式の分布を見い出す場合に，よく適用される．一般に，k 変数正規分布 (7.4.1) における 2 次形式，すなわち，(7.4.2) で定義される $Q(x_1,\cdots,x_k)$ がガンマ分布に従うことを示そう．

(7.5.7) $$\varphi(t)=\frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}}\int_{R_k}e^{-\frac{1}{2}(1-it)Q(x_1,\cdots,x_k)}\,dx_1\cdots dx_k$$

で定義される $\frac{1}{2}Q(x_1,\cdots,x_k)$ の特性関数について考える．

変換 $y_i=\sqrt{1-it}\,(x_i-\mu_i)$，$i=1,\cdots,k$ を行なうと，(7.5.7) は

(7.5.8) $$\varphi(t)=\frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}}(1-it)^{-\frac{1}{2}k}\int_{R_k}\exp\left(-\frac{1}{2}\sum_{i,j=1}^{k}\sigma^{ij}y_iy_j\right)dy_1\cdots dy_k$$

となる．(7.4.3) が使えて

(7.5.9) $$\varphi(t)=(1-it)^{-\frac{1}{2}k}.$$

これはガンマ分布 $G(k/2)$ の特性関数に他ならない．

7.5.2 $(x_1,\cdots,x_k)$ が k 次元分布 $N(\{\mu_i\},\|\sigma_{ij}\|)$，$i,j=1,\cdots,k$ に従うベクトル確率変数のとき，$\frac{1}{2}Q(x_1,\cdots,x_k)$ はガンマ分布 $G(k/2)$ を持つ．

連続待ち時間分布もガンマ分布に関連している．6.5 節より，(6.5.14) は，試行を x 回行なうまで（k 個の C を得るまで），待たねばならない確率を表わしている．ここで

は，C を連続時間にわたって起り得る事象とし，k 個の C を得るまで何時間待たねばならないかを考えよう．$(0,t)$ を，ちょうど $(k-1)$ 個の C が起る時間区間，$(t, t+\Delta t)$ を k 番目の C が起る時間区間とする．区間 $(0,t)$ を長さ Δt の区間に $t/\Delta t$ 等分する．長さ Δt の個々の区間で C が起る確率を $\lambda\Delta t + o(\Delta t)$ とする．ただし，λ は定数で，単位時間当りの C の平均生起回数である．さらに，時間区間 I での C の生起は，I と重復しない区間 I' での生起と独立であるとする．ちょうど k 個の C が起るまで（時間区間 $(t,t+\Delta t)$ まで）待ち続けなければならない確率を求めたい．これは，$t/\Delta t$ 等分された区間 $(0,t)$ で $(k-1)$ 個の C が起る確率と，区間 $(t,t+\Delta t)$ で k 番目の C が起る確率の積に等しい．$(\Delta t)^2$ およびこれ以上の次数を無視して $x=t/\Delta t$，$p=\lambda\Delta t$ と置く．(6.5.15) より，この確率は

(7.5.10)
$$\binom{\frac{t}{\Delta t}}{k-1}(\lambda\Delta t)^k(1-\lambda\Delta t)^{(t/\Delta t)-k+1}$$
$$= t\,\frac{(t-\Delta t)(t-2\Delta t)\cdots(t-(k-2)\Delta t)}{(k-1)!}\lambda^{k-1}(1-\lambda\Delta t)^{(t/\Delta t)-k+1}\lambda\Delta t.$$

Δt で割って，$\Delta t \longrightarrow 0$ とすると，p. d. f.

(7.5.11)
$$f(t)=\begin{cases}\dfrac{\lambda(\lambda t)^{k-1}e^{-\lambda t}}{(k-1)!}, & t>0\\ 0, & \text{その他}\end{cases}$$

を得る．これは，ここで用いた仮定のもとで，ちょうど k 個の C を得るまでに要した待ち時間の p. d. f. である．(7.5.11) から，λt がガンマ分布 $G(k)$ を持つ確率変数であることは明らかである．

最後に，次の結果を述べておく．これは読者自身で確かめてみよ．

7.5.3　x が矩形分布 $R\left(\frac{1}{2},1\right)$ に従う確率変数のとき，確率変数 $y=-\log x$ はガンマ分布 $G(1)$ を持つ．

7.6 ベータ分布

(7.6.1)
$$f(x)=\begin{cases}\dfrac{\Gamma(\nu_1+\nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)}x^{\nu_1-1}(1-x)^{\nu_2-1}, & 0<x<1\\ 0, & \text{その他}\end{cases}$$

で定義される p. d. f. を持つ分布も統計学によく現われる連続分布である．ただし，ν_1，ν_2 は正の実数．p. d. f. (7.6.1) を持つ分布をベータ分布 $Be(\nu_1,\nu_2)$ という．

$f(x)$ の $(0,1)$ 上での積分が 1 であることを見るには，次のようにする．(7.5.1) を用いて

(7.6.2)
$$\Gamma(\nu_1)\Gamma(\nu_2)=\int_0^\infty\int_0^\infty x_1^{\nu_1-1}x_2^{\nu_2-1}e^{-x_1-x_2}\,dx_1\,dx_2.$$

変換
$$x_1=r^2\cos^2\theta,\qquad x_2=r^2\sin^2\theta$$
のヤコビアンは
$$\left|\frac{\partial(x_1,x_2)}{\partial(r,\theta)}\right|=4r^3\sin\theta\cos\theta$$
だから

(7.6.3)
$$\Gamma(\nu_1)\Gamma(\nu_2)=4\int_0^{\frac{1}{2}\pi}\int_0^\infty(\cos\theta)^{2\nu_1-1}(\sin\theta)^{2\nu_2-1}r^{2\nu_1+2\nu_2-1}e^{-r^2}\,d\theta\,dr.$$

さらに，変換 $r=\sqrt{y}$ を
$$2\int_0^\infty r^{2\nu_1+2\nu_2-1}e^{-r^2}\,dr$$
に施すと
$$\Gamma(\nu_1+\nu_2)=2\int_0^\infty r^{2\nu_1+2\nu_2-1}e^{-r^2}\,dr.$$

(7.6.3) は

(7.6.4)
$$\frac{\Gamma(\nu_1)\Gamma(\nu_2)}{\Gamma(\nu_1+\nu_2)}=2\int_0^{\frac{1}{2}\pi}(\cos\theta)^{2\nu_1-1}(\sin\theta)^{2\nu_2-1}\,d\theta.$$

最後に，$\cos\theta=\sqrt{x}$ を (7.6.4) に適用すると

(7.6.5)
$$\frac{\Gamma(\nu_1)\Gamma(\nu_2)}{\Gamma(\nu_1+\nu_2)}=\int_0^1 x^{\nu_1-1}(1-x)^{\nu_2-1}\,dx$$

になる．ゆえに，(7.6.1) の $f(x)$ の $(0,1)$ における積分は 1 になる．

ベータ分布 $Be(\nu_1,\nu_2)$ の c. d. f. $F(x)$ を $I_x(\nu_1,\nu_2)$ で表わし，これを不完全ベータ関数という．これに関しては，Karl Pearson (1934) は x が 0.01 から 1.00，ν_1,ν_2 が 0.05 から 50 までを *The Tables of the Incomplete Beta Function* にまとめている．(7.6.5) の右辺の定積分で定義される ν_1，ν_2 の関数をベータ関数と呼び，$B(\nu_1,\nu_2)$ で表わす．

ベータ分布の特性関数は，モーメントを求めるためには有用ではない．(7.6.1) のモー

メントは直接計算される．μ_r' を r 次のモーメントとすると

(7.6.6) $$\mu_r' = \frac{\Gamma(\nu_1+\nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)}\int_0^1 x^{\nu_1+r-1}(1-x)^{\nu_2-1}\,dx.$$

(7.6.5) を用いて

(7.6.7) $$\mu_r' = \frac{\Gamma(\nu_1+\nu_2)\Gamma(\nu_1+r)}{\Gamma(\nu_1+\nu_2+r)\Gamma(\nu_1)}.$$

したがって

(7.6.8) $$\mu(x) = \frac{\nu_1}{\nu_1+\nu_2}, \qquad \sigma^2(x) = \frac{\nu_1\nu_2}{(\nu_1+\nu_2)^2(\nu_1+\nu_2+1)}.$$

数理統計学では，ベータ分布は順序統計量の理論や統計的検定によく現われてくる．これに関しては，後章で議論する．次に，最も重要な場合を述べよう．

7.6.1 x_1，x_2 をそれぞれ独立にガンマ分布 $G(\nu_1)$，$G(\nu_2)$ に従う確率変数とする．このとき，確率変数

(7.6.9) $$u = \frac{x_1}{x_1+x_2}$$

はベータ分布 $Be(\nu_1,\nu_2)$ を持つ．

これは容易に確かめられる．x_1 と x_2 の p.e. は

(7.6.10) $$\frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)}x_1^{\nu_1-1}x_2^{\nu_2-1}e^{-x_1-x_2}\,dx_1\,dx_2.$$

これに変換

(7.6.11) $$u = \frac{x_1}{x_1+x_2}, \qquad v = x_2$$

を行なうと，u と v の p.e. は

(7.6.12) $$\frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)}u^{\nu_1-1}v^{\nu_1+\nu_2-1}(1-u)^{-(\nu_1+1)}e^{-v/(1-u)}\,du\,dv.$$

u の分布は p.e. (7.6.12) を持つ分布の周辺分布であるから，(7.6.12) を v について，0 から ∞ まで積分すると，u の分布が $Be(\nu_1,\nu_2)$ として得られる．

注意　ベータ分布を終える前に，ガンマ分布に関する重要な公式を 2 つ述べておこう．これはルジャンドルの公式とスターリングの公式であるが，ベータ分布を少し変形すれば得られる．

ガンマ関数に対するルジャンドルの重複公式

(7.6.5) において，$\nu_1=\nu_2=\nu$ とすると

(7.6.13) $$\frac{\Gamma^2(\nu)}{\Gamma(2\nu)} = \int_0^1 x^{\nu-1}(1-x)^{\nu-1}\,dx = 2\int_0^{\frac{1}{2}}(x-x^2)^{\nu-1}\,dx.$$

変数変換 $y=4(x-x^2)$，$0<x<\dfrac{1}{2}$，すなわち $x=(1-\sqrt{1-y})/2$ を行なうと

(7.6.14) $$\frac{\Gamma^2(\nu)}{\Gamma(2\nu)} = 2^{1-2\nu}\int_0^1 y^{\nu-1}(1-y)^{-\frac{1}{2}}\,dy = 2^{1-2\nu}\frac{\Gamma(\nu)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\nu+\frac{1}{2}\right)}.$$

$\Gamma\left(\dfrac{1}{2}\right)=\sqrt{\pi}$ に注意すれば，ガンマ関数に対するルジャンドルの重複公式：

(7.6.15) $$\Gamma(2\nu) = \frac{2^{2\nu-1}}{\sqrt{\pi}}\Gamma(\nu)\Gamma\left(\nu+\frac{1}{2}\right)$$

を得る．これはあとで用いられる．

大きな階乗に対するスターリングの公式

ベータ分布のある種の極限性質を用いると，簡単な計算によって，g が十分大きいとき，$\Gamma(g)$ に対する高精度の近似式が得られる．

(7.6.1) で，$\nu_1=g$，$\nu_2=n+1$ とし，変数変換 $x=y/n$ を行なうと，

(7.6.16) $$\frac{\Gamma(g+n+1)}{n^g\Gamma(g)\Gamma(n+1)}\int_0^n y^{g-1}\left(1-\frac{y}{n}\right)^n dy = 1$$

がすべての n について成り立つ．ゆえに，$n\longrightarrow\infty$ のとき，(7.6.16) の極限は 1 になる．(7.6.16) の被積分関数は，$n\longrightarrow\infty$ のとき，任意の有限区間 $(0,K)$ 上で一様に $y^{g-1}e^{-y}$ に収束する．n を十分大きくすると，差

$$\int_0^{\min(n,K)} y^{g-1}\left(1-\frac{y}{n}\right)^n dy - \int_0^K y^{g-1}e^{-y}\,dy$$

はいくらでも 0 に近づけることができる．しかし，最初から

$$\int_K^\infty y^{g-1}e^{-y}\,dy$$

をいくらでも小さくなるように K が選べる．ゆえに

(7.6.17) $$\lim_{n\to\infty}\int_0^n y^{g-1}\left(1-\frac{y}{n}\right)^n dy = \int_0^\infty y^{g-1}e^{-y}\,dy = \Gamma(g).$$

よって
$$\lim_{n\to\infty}\frac{\Gamma(n+1)n^g}{\Gamma(g+n+1)} = 1.$$

これは次のように書ける．

(7.6.18) $$\Gamma(g) = \lim_{n\to\infty}\frac{n!n^g}{g(g+1)\cdots(g+n)}.$$

さて

(7.6.19) $$S_n(g) = g\log n + \sum_{\alpha=1}^n \log\alpha - \sum_{\alpha=0}^n \log(g+\alpha)$$

を g で微分すると

(7.6.20) $$S_n'(g) = \log n - \sum_{\alpha=0}^n \frac{1}{g+\alpha}.$$

よって

$$\lim_{n\to\infty} S_n(g) = \log\Gamma(g), \qquad \lim_{n\to\infty} S'_n(g) = \frac{\Gamma'(g)}{\Gamma(g)}.$$

簡単な計算により，次の評価式を得る．

(7.6.21) $$-B_n(g) < -\sum_{\alpha=0}^{n}\frac{1}{g+\alpha} + \frac{1}{2g} + \frac{1}{2(g+n)} < -C_n(g).$$

ただし

$$B_n(g) = \frac{1}{2}\int_{\frac{1}{2}}^{n+\frac{1}{2}}\left(\frac{1}{g+x-1} + \frac{1}{g+x}\right)dx, \qquad C_n(g) = \int_{\frac{1}{2}}^{n+\frac{1}{2}}\frac{dx}{g+x-\frac{1}{2}}.$$

$\log n - \frac{1}{2}(1/g + 1/(g+n))$ を $A_n(g)$ とし，これを (7.6.21) の両辺に加えると

(7.6.22) $$A_n(g) - B_n(g) < S'_n(g) < A_n(g) - C_n(g).$$

$A_n(g)$, $B_n(g)$, $C_n(g)$ を評価し，(7.6.21) で $n \longrightarrow \infty$ とすると

(7.6.23) $$\log g - \frac{1}{2g} + \frac{1}{2}\log\left(1 - \frac{1}{4g^2}\right) < \frac{\Gamma'(g)}{\Gamma(g)} < \log g - \frac{1}{2g}.$$

したがって

(7.6.24) $$\frac{\Gamma'(g)}{\Gamma(g)} = \log g - \frac{1}{2g} + O\left(\frac{1}{g^2}\right).$$

g で積分すると

(7.6.25) $$\log\Gamma(g) = g\log g - g - \frac{1}{2}\log g + C + O\left(\frac{1}{g}\right).$$

ただし，C は g に無関係な定数であるが，(7.6.15) を用いると，この C の値が定まる．ν を g で置き換えて，対数をとると

(7.6.26) $$\log\Gamma(2g) = (2g-1)\log 2 + \log\Gamma(g) + \log\Gamma\left(g + \frac{1}{2}\right) - \frac{1}{2}\log\pi.$$

(7.6.25) で与えられる式を (7.6.26) に代入して，$g \longrightarrow \infty$ とすれば，$C = \frac{1}{2}\log 2\pi$ となる．したがって

(7.6.27) $$\Gamma(g) = \sqrt{2\pi}\, g^{g-\frac{1}{2}}e^{-g}\left(1 + O\left(\frac{1}{g}\right)\right)$$

を得る．

不等式 (7.6.21) は，(7.6.27) で $O(1/g)$ における $1/g$ の係数を定めるほど厳密ではない．事実，$O(1/g) = 1/12g + 1/288g^2 + O(1/g^3)$ であるが，これを示すには，ここで用いられた方法よりも，より厳密な評価式が必要である．たとえば，Whittaker と Watson (1927) や Cramér (1946) に詳しい．十分大きな g については，漸近公式

(7.6.28) $$\Gamma(g) \cong \sqrt{2\pi}\, g^{g-\frac{1}{2}}e^{-g}$$

がよく用いられる．g が正整数のとき，$\Gamma(g+1) = g!$ だから，大きな階乗に対するスターリングの公式：

(7.6.29) $$g! \cong \sqrt{2\pi g}\, g^g e^{-g}$$

が得られる．

7.7　ディリクレ分布

(7.6.1) の類似として，k 変数の場合は p.d.f.

(7.7.1) $$\begin{aligned} f(x_1,\cdots,x_k) &= \frac{\Gamma(\nu_1+\cdots+\nu_{k+1})}{\Gamma(\nu_1)\cdots\Gamma(\nu_{k+1})}x_1^{\nu_1-1}\cdots x_k^{\nu_k-1}(1-x_1-\cdots-x_k)^{\nu_{k+1}-1}, \\ &\qquad \text{単体 } S_k : \left\{\begin{array}{c}(x_1,\cdots,x_k) : x_i \geqslant 0 \\ i = 1,\cdots,k \\ \sum_{i=1}^{k} x_i \leqslant 1\end{array}\right\} \\ &= 0, \qquad \text{その他} \end{aligned}$$

を持つ分布がある．ここで ν_i はすべて正の実数である．p.d.f. (7.7.1) を持つ分布を k 変数**ディリクレ分布** $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ という．$k=1$ のとき，$D(\nu_1;\nu_2)$ は $Be(\nu_1,\nu_2)$ に一致する．すなわち，(7.7.1) はベータ分布 $Be(\nu_1,\nu_2)$ の p.d.f. に帰着する．8.7 節で見るように，順序統計量の計算にはディリクレ分布が基礎になっている．

$f(x_1,\cdots,x_k)$ の単体 S_k 上の積分が 1 であることは，次によって確かめられる．変換

$$\begin{aligned} x_1 &= \theta_1 \\ x_2 &= \theta_2(1-x_1) \\ &\ \ \vdots \\ x_k &= \theta_k(1-x_1-\cdots-x_{k-1}) \end{aligned}$$

を (7.7.1) に適用すると

(7.7.2) $$\begin{aligned} f(x_1,\cdots,x_k)\,dx_1\cdots dx_k &= \frac{\Gamma(\nu_1+\cdots+\nu_{k+1})}{\Gamma(\nu_1)\cdots\Gamma(\nu_{k+1})}\theta_1^{\nu_1-1}(1-\theta_1)^{\nu_2+\cdots+\nu_{k+1}-1}\theta_2^{\nu_2-1} \\ &\quad \cdot(1-\theta_2)^{\nu_3+\cdots+\nu_{k+1}-1}\cdots \\ &\quad \cdot\theta_k^{\nu_k-1}(1-\theta_k)^{\nu_{k+1}-1}\,d\theta_1\cdots d\theta_k. \end{aligned}$$

ここに，θ の範囲は k 次元単位立方体 $\{(\theta_1,\cdots,\theta_k) : 0 \leqslant \theta_i \leqslant 1,\ i = 1,\cdots,k\}$ である．ここで，(7.6.5) が使えて

(7.7.3)
$$\int_{S_k} f(x_1,\cdots,x_k)\,dx_1\cdots dx_k$$
$$=\frac{\Gamma(\nu_1+\cdots+\nu_{k+1})}{\Gamma(\nu_1)\cdots\Gamma(\nu_{k+1})}\cdot\frac{\Gamma(\nu_1)\Gamma(\nu_2+\cdots+\nu_{k+1})}{\Gamma(\nu_1+\cdots+\nu_{k+1})}\cdots\frac{\Gamma(\nu_k)\Gamma(\nu_{k+1})}{\Gamma(\nu_k+\nu_{k+1})}.$$

よって，右辺は 1 になる．

したがって，積分

(7.7.4)
$$\int_{S_k} x_1^{\nu_1-1}\cdots x_k^{\nu_k-1}(1-x_1-\cdots-x_k)^{\nu_{k+1}-1}\,dx_1\cdots dx_k$$

の値は

(7.7.5)
$$\frac{\Gamma(\nu_1)\cdots\Gamma(\nu_{k+1})}{\Gamma(\nu_1+\cdots+\nu_{k+1})}$$

になる．この積分は Dirichlet (1839) が最初に研究したので，**ディリクレ積分**と呼ばれている．(7.7.4)，(7.7.5) がベータ分布における (7.6.5) に対応していることを考慮すれば，(7.7.1) を p.d.f. に持つ分布をディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ と呼ぶのは，(7.6.1) を p.d.f. に持つ分布をベータ分布と呼ぶことと類似している．

k 変数ディリクレ分布のモーメントは

(7.7.6)
$$\mu'_{r_1\cdots r_k}=\frac{\Gamma(\nu_1+r_1)\cdots\Gamma(\nu_k+r_k)\Gamma(\nu_1+\cdots+\nu_{k+1})}{\Gamma(\nu_1+\cdots+\nu_{k+1}+r_1+\cdots+r_k)\Gamma(\nu_1)\cdots\Gamma(\nu_k)}$$

になることが容易に確かめられる．よって，x の平均，分散，共分散は

$$\mu(x_i)=\frac{\nu_i}{\nu_1+\cdots+\nu_{k+1}},\qquad i=1,\cdots,k$$

(7.7.7)
$$\sigma^2(x_i)=\frac{\nu_i(\nu_1+\cdots+\nu_{k+1}-\nu_i)}{(\nu_1+\cdots+\nu_{k+1})^2(\nu_1+\cdots+\nu_{k+1}+1)},\qquad i=1,\cdots,k$$

$$\sigma(x_i,x_j)=-\frac{\nu_i\nu_j}{(\nu_1+\cdots+\nu_{k+1})^2(\nu_1+\cdots+\nu_{k+1}+1)},\ i\neq j=1,\cdots,k.$$

7.6.1 は k 変数の場合には次のようになる．

7.7.1　$x_1,\cdots,x_{k+1}$ をガンマ分布 $G(\nu_1),\cdots,G(\nu_{k+1})$ に従う独立な確率変数とする．このとき

(7.7.8)
$$y_i=\frac{x_i}{x_1+\cdots+x_{k+1}},\qquad i=1,\cdots,k$$

とすれば，$(y_1,\cdots,y_k)$ は k 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ を持つ．

7.7.1 の証明は簡単である．読者自身で証明せよ．$k_1<k$ として，(7.7.6) のモーメントで $r_{k_1+1}=\cdots=r_k=0$ とする．$(x_1,\cdots,x_k)$ が k 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ に従うとすれば，$(x_1,\cdots,x_{k_1})$ の周辺分布のモーメント $\mu'_{r_1\cdots r_{k_1}}$ が

(7.7.9)
$$\mu'_{r_1\cdots r_{k_1}}=\frac{\Gamma(\nu_1+r_1)\cdots\Gamma(\nu_{k_1}+r_{k_1})\Gamma(\nu_1+\cdots+\nu_{k+1})}{\Gamma(\nu_1+\cdots+\nu_{k+1}+r_1+\cdots+r_{k_1})\Gamma(\nu_1)\cdots\Gamma(\nu_{k_1})}$$

として得られる．しかし，これは k_1 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_{k_1};\nu_{k_1+1}+\cdots+\nu_{k+1})$ のモーメントである．**5.5.1 a** を多次元の場合に適用すれば，分布は一意に定まるから，次が成り立つ．

7.7.2　$(x_1,\cdots,x_k)$ を k 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ に従う確率変数とすると，$(x_1,\cdots,x_{k_1})$，$k_1<k$ の周辺分布は k_1 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_{k_1};\nu_{k_1+1}+\cdots+\nu_{k+1})$ になる．

k 変数ディリクレ分布に対する条件つき確率変数 $x_k|x_1,\cdots,x_{k-1}$ の p.d.f. は $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ の p.d.f. と $D(\nu_1,\cdots,\nu_{k-1};\nu_k+\nu_{k+1})$ の p.d.f. の比として書き表わせる．したがって，$x_k|x_1,\cdots,x_{k-1}$ の p.e. は次のようになる．

(7.7.10)
$$dF(x_k|x_1,\cdots,x_{k-1})$$
$$=\frac{\Gamma(\nu_k+\nu_{k+1})}{\Gamma(\nu_k)\Gamma(\nu_{k+1})}\left(\frac{x_k}{1-x_1-\cdots-x_{k-1}}\right)^{\nu_k-1}$$
$$\cdot\left(1-\frac{x_k}{1-x_1-\cdots-x_{k-1}}\right)^{\nu_{k+1}-1}d\left(\frac{x_k}{1-x_1-\cdots-x_{k-1}}\right).$$

(7.7.10) より，明らかに次が成り立つ．

7.7.3　$(x_1,\cdots,x_k)$ が k 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ に従うとき，条件つき確率変数 $x_k|x_1,\cdots,x_{k-1}$ は次の性質を持つ．すなわち

$$\left.\frac{x_k}{1-x_1-\cdots-x_{k-1}}\right|x_1,\cdots,x_{k-1}$$

はベータ分布 $Be(\nu_k,\nu_{k+1})$ に従う．

(7.6.8) を参考にすると，$x_k|x_1,\cdots,x_{k-1}$ の平均と分散は

$$\mu(x_k|x_1,\cdots,x_{k-1})=\frac{\nu_k}{\nu_k+\nu_{k+1}}(1-x_1-\cdots-x_{k-1})$$

(7.7.11)
$$\sigma^2(x_k|x_1,\cdots,x_{k-1})$$
$$=\frac{\nu_k\nu_{k+1}}{(\nu_k+\nu_{k+1})^2(\nu_k+\nu_{k+1}+1)}(1-x_1-\cdots-x_{k-1})^2.$$

k 変数ディリクレ分布は，他に，次の性質を持つ．

7.7.4　$(x_1,\cdots,x_k)$ が k 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ に従うとき，和

$x_1+\cdots+x_k$ はベータ分布 $Be(\nu_1+\cdots+\nu_k,\ \nu_{k+1})$ を持つ.

これは次のようにして証明できる．$1-(x_1+\cdots+x_k)$ の r 次階乗モーメントは

$$\frac{\Gamma(\nu_{k+1}+r)\Gamma(\nu_1+\cdots+\nu_{k+1})}{\Gamma(\nu_1+\cdots+\nu_{k+1}+r)\Gamma(\nu_{k+1})}, \qquad r=1,2,\cdots.$$

したがって，**5.5.1 a** より，$1-(x_1+\cdots+x_k)$ はベータ分布 $Be(\nu_{k+1},\nu_1+\cdots+\nu_k)$ に従う．ゆえに $x_1+\cdots+x_k$ はベータ分布 $Be(\nu_1+\cdots+\nu_k,\nu_{k+1})$ を持つ．

一般には，**7.7.4** は次のように拡張される．

7.7.5　$(x_1,\cdots,x_k)$ が k 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ に従うとき，確率変数 $(z_1,\cdots,z_s)$ は s 変数ディリクレ分布 $D(\nu_{(1)},\cdots,\nu_{(s)};\nu_{(s+1)})$ を持つ．ただし $z_1=x_1+\cdots+x_{k_1}$，$z_2=x_{k_1+1}+\cdots+x_{k_1+k_2}$，$\cdots$，$z_s=x_{k_1+\cdots+k_{s-1}+1}+\cdots+x_{k_1+\cdots+k_s}$ で，$k_1+\cdots+k_s\leqslant k$，$\nu_{(1)}=\nu_1+\cdots+\nu_{k_1}$，$\cdots$，$\nu_{(s)}=\nu_{k_1+\cdots+k_{s-1}+1}+\cdots+\nu_{k_1+\cdots+k_s}$，$\nu_{(s+1)}=\nu_{k_1+\cdots+k_s+1}+\cdots+\nu_{k+1}$.

まず，(7.7.1) から

(7.7.12)
$$\mathscr{E}[x_{k_1+1}^{r_{k_1+1}}\cdots x_k^{r_k}(1-z_1-x_{k_1+1}-\cdots-x_k)^{r_{(1)}}]$$
$$=G\cdot\frac{\Gamma(\nu_{(1)})\Gamma(\nu_{k_1+1}+r_{k_1+1})\cdots\Gamma(\nu_k+r_k)\Gamma(\nu_{k+1}+r_{(1)})}{\Gamma(\nu_{(1)}+\nu_{k_1+1}+\cdots+\nu_k+\nu_{k+1}+r_{k_1+1}+\cdots+r_k+r_{(1)})}.$$

ただし

(7.7.13)
$$G=\frac{\Gamma(\nu_{(1)}+\nu_{k_1+1}+\cdots+\nu_{k+1})}{\Gamma(\nu_{(1)})\Gamma(\nu_{k_1+1})\cdots\Gamma(\nu_{k+1})}.$$

しかし，(7.7.12) の右辺は p.d.f.

(7.7.14)
$$f(z_1,x_{k_1+1},\cdots,x_k)$$
$$=Gz_1^{\nu_{(1)}-1}x_{k_1+1}^{\nu_{k_1+1}-1}\cdots x_k^{\nu_k-1}(1-z_1-x_{k_1+1}-\cdots-x_k)^{\nu_{k+1}-1}$$

から計算される

$$\mathscr{E}[x_{k_1+1}^{r_{k_1+1}}\cdots x_k^{r_k}(1-z_1-x_{k_1+1}-\cdots-x_k)^{r_{(1)}}]$$

の値である．$(z_1,x_{k_1+1},\cdots,x_k)$ は有界な確率変数だから，(7.7.14) の p.d.f. は一意に定まる．

この方法を $z_2,z_3,\cdots$ に適用すれば

(7.7.15)
$$f(z_1,\cdots,z_s)$$
$$=\frac{\Gamma(\nu_{(1)}+\cdots+\nu_{(s+1)})}{\Gamma(\nu_{(1)})\cdots\Gamma(\nu_{(s+1)})}z_1^{\nu_{(1)}-1}\cdots z_s^{\nu_{(s)}-1}(1-z_1-\cdots-z_s)^{\nu_{(s+1)}-1}$$

が得られ，**7.7.5** が成立する．

8.7 節で順序統計量を取り扱う場合，ここで k 変数ディリクレ分布に関連する分布を紹介しておくと役立つであろう．$(x_1,\cdots,x_k)$ が k 変数ディリクレ分布 $D(\nu_1,\cdots,\nu_k;\nu_{k+1})$ を持つとき

(7.7.16)
$$\begin{array}{l} y_1=x_1\\ y_2=x_1+x_2\\ \quad\vdots\\ y_k=x_1+\cdots+x_k \end{array}$$

とする．この変換のヤコビアンは 1 だから，$(y_1,\cdots,y_k)$ の p.d.f. は

(7.7.17)
$$f(y_1,\cdots,y_k)=\frac{\Gamma(\nu_1+\cdots+\nu_{k+1})}{\Gamma(\nu_1)\cdots\Gamma(\nu_{k+1})}y_1^{\nu_1-1}$$
$$\cdot(y_2-y_1)^{\nu_2-1}\cdots(y_k-y_{k-1})^{\nu_k-1}(1-y_k)^{\nu_{k+1}-1}.$$

ただし，y の範囲は $0<y_1<\cdots<y_k<1$．(7.7.17) を p.d.f. として持つ分布を順序づけられた k 変数ディリクレ分布と呼び，$D^*(\nu_1,\cdots,\nu_k;\nu_{k+1})$ と書く．$k=1$ のとき，(7.7.17) はベータ分布 $Be(\nu_1,\nu_2)$ の p.d.f. になる．

この分布に関連して，s 個の y の部分集合からなる確率変数の周辺分布が，順序統計量に関する問題によく現われる．これを述べよう．

7.7.6　$(y_1,\cdots,y_k)$ が順序づけられた k 変数ディリクレ分布 $D^*(\nu_1,\cdots,\nu_k;\nu_{k+1})$ を持つとき，$(y_{k_1},y_{k_1+k_2},\cdots,y_{k_1+\cdots+k_s})$ の周辺分布は順序づけられた s 変数ディリクレ分布 $D^*(\nu_{(1)},\cdots,\nu_{(s)};\nu_{(s+1)})$ になる．ただし，$\nu_{(1)},\cdots,\nu_{(s+1)}$ は **7.7.5** で定義されている．

(7.7.16) で与えられた確率変数 $y_{k_1},\cdots,y_{k_1+\cdots+k_s}$ は，**7.7.5** で定義された $z_1,z_1+z_2,\cdots,z_1+\cdots+z_s$ と分布が同じであることに注意せよ．なぜなら，$(z_1,\cdots,z_s)$ は s 変数ディリクレ分布 $D(\nu_{(1)},\cdots,\nu_{(s)};\nu_{(s+1)})$ に従うから，定義より，$(z_1,z_1+z_2,\cdots,z_1+\cdots+z_s)$ は順序づけられた s 変数ディリクレ分布 $D^*(\nu_{(1)},\cdots,\nu_{(s)};\nu_{(s+1)})$ を持つ．

7.7.6 は (7.7.17) を $y_{k_1},y_{k_1+k_2},\cdots,y_{k_1+\cdots+k_s}$ 以外の y で直接積分しても得られる．すなわち，まず $0<y_1<\cdots<y_{k_1}$ なる範囲を $y_1,\cdots,y_{k_1-1}$ で積分し，次に，$y_{k_1}<y_{k_1+1}<\cdots<y_{k_1+k_2}$ なる範囲を $y_{k_1+1},\cdots,y_{k_1+k_2-1}$ で積分する．以下，これを繰り返せばよい．

7.8　分散分析に関する分布

カイ 2 乗分布，スチューデント分布，スネディッカー分布の 3 つの分布と，ベータ分布，ガンマ分布とは密接に関係がある．この 3 つの分布は，分散分析を行なう場合や，後章で議論されているように，正規分布に従う確率変数を統計的に処理する場合にも，基本的な役割を果たすであろう．ここでは，分布の基本的な型を，重要な連続型分布の例題として述べておくと便利であろう．

(a)　カイ 2 乗分布

ガンマ分布 $G(\mu)$ の p.e. に変数変換

$$x = \chi^2/2$$

を施すと

(7.8.1)
$$dF_{2\mu}(\chi^2) = \frac{\left(\dfrac{\chi^2}{2}\right)^{\mu-1} e^{-\frac{1}{2}\chi^2}}{2\Gamma(\mu)} d(\chi^2).$$

これが自由度 2μ のカイ 2 乗分布の p.e. である．この分布を持つ確率変数を簡単にカイ 2 乗分布 $C(2\mu)$ を持つという．この分布の統計的な適用例においては，多くの場合，2μ は正の整数である．$\int_{\chi^2_\alpha}^{\infty} dF_{2\mu}(\chi^2) = \alpha$ となる値 χ^2_α については，α が 0.001 から 0.99，2μ が $1, 2, \cdots, 30$ まで Fisher (1925 *a*) が表にしている．さらに詳しい表は E. S. Pearson と Hartley (1954) 編 *Biometrika Tables for Statisticians* にまとめられている．

カイ 2 乗分布 $C(2\mu)$ の特性関数は，(7.5.6) で t を $2t$ に置き換えると

(7.8.2)
$$\varphi(t) = (1-2it)^{-\mu}$$

として得られる．(7.8.2) の型から

7.8.1　カイ 2 乗分布 $C(2\mu)$ は μ に関して再生的である．

$\chi^2/2$ はガンマ分布 $G(\mu)$ を持つから，(7.5.3) より，χ^2 の r 次モーメントは

$$\mu_r' = \frac{2^r\Gamma(\mu+r)}{\Gamma(\mu)}.$$

したがって，$C(2\mu)$ の平均と分散は

$$\mathscr{E}(\chi^2) = 2\mu, \qquad \sigma^2(\chi^2) = 4\mu.$$

カイ 2 乗の概念は K. Pearson (1900) により紹介されたが，統計に関する文献に頻繁に現われるようになってから，その分布がよく用いられるようになった．

7.5.2 のガンマ分布を用いると，$Q(x_1, \cdots, x_k)$ の分布は，次のようにカイ 2 乗の概念を用いて表わされる．

7.8.2　$(x_1, \cdots, x_k)$ が k 変数正規分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$, $i, j = 1, \cdots, k$ に従うとき，$\sum_{i,j=1}^{k} \sigma^{ij}(x_i-\mu_i)(x_j-\mu_j)$ はカイ 2 乗分布 $C(k)$ を持つ．

この系として

7.8.2 a　x が正規分布 $N(0,1)$ に従うならば，x^2 はカイ 2 乗分布 $C(1)$ に従う．

(7.8.1) において自由度が 2μ であるというのは，**7.8.2** での自由度が，非退化正規分布に含まれる確率変数の個数に等しいことから明らかである．

このように，自由度の考え方は，後の章で，カイ 2 乗分布を議論して行くにつれて，より明確になるであろう．

(b)　スチューデント分布

この分布の本質とこれがどの確率変数に由来しているかは，以下に述べられている．

7.8.3　u が分布 $N(0,1)$ を持ち，v がカイ 2 乗分布 $C(k)$ に従うとする．u と v が独立のとき，確率変数

(7.8.3)
$$t = \frac{u}{\sqrt{v/k}}$$

の p.d.f は

(7.8.4)
$$f_k(t) = \frac{\Gamma\left(\dfrac{k+1}{2}\right)}{\Gamma\left(\dfrac{k}{2}\right)\sqrt{\pi k}}\left(1+\frac{t^2}{k}\right)^{-\frac{1}{2}(k+1)}$$

になる．

(7.8.4) を自由度 k のスチューデント分布[*]，あるいは，簡単にスチューデント分布 $S(k)$[**] と呼ぶ．8.4 節および 10.4 節では，この分布がしばしば適用されている．

(7.8.4) を示そう．変換

[*] スチューデントの t 分布，あるいは単に t 分布ともいわれている．(訳注)
[**] 自由度 k の t 分布に対応して，$t(k)$ と書くこともある．(訳注)

(7.8.5)
$$s=v,\qquad t=\frac{u}{\sqrt{v/k}}$$

を，u と v の p.e.

(7.8.6)
$$\frac{1}{2\sqrt{2\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(\frac{v}{2}\right)^{\frac{1}{2}(k-2)}e^{-\frac{1}{2}(u^2+v)}\,du\,dv$$

に施して，t の周辺分布をとる．

変換 (7.8.5) のヤコビアンは $\sqrt{s/k}$ だから，s と t の p.e. は

(7.8.7)
$$\frac{1}{2\sqrt{\pi k}\,\Gamma\left(\frac{k}{2}\right)}\left(\frac{s}{2}\right)^{\frac{1}{2}(k-1)}e^{-\frac{1}{2}[1+(t^2/k)]s}\,ds\,dt.$$

t の周辺分布をとる．(すなわち，s で 0 から ∞ まで積分すると) t の p.d.f. は (7.8.4) で与えられることがわかる．

$P(|t|>t_\alpha)=\alpha$ となる t_α については，Fisher (1925 *a*) が，α が 0.1 から 0.99，k が 1, 2, …, 30，まで表にしている．また Pearson と Hartley (1945) の *Biometrika Tables for Statisticians* にも記載されている．

p.d.f. (7.8.4) を持つ分布の奇数次のモーメントはすべて 0 である．偶数次のモーメントは存在すれば

$$\mu'_{2r}=\mu_{2r}=k^r\mathscr{E}\left(\frac{u^{2r}}{v^r}\right).$$

u と v が独立だから

$$\mathscr{E}\left(\frac{u^{2r}}{v^r}\right)=\mathscr{E}(u^{2r})\cdot\mathscr{E}(v^{-r}).$$

しかし，u^2，v は独立で，それぞれカイ 2 乗分布 $C(1)$，$C(k)$ に従うから

(7.8.8)
$$\mu'_{2r}=\mu_{2r}=k^r\frac{\Gamma\left(\frac{1}{2}+r\right)\Gamma\left(\frac{k}{2}-r\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{k}{2}\right)}.$$

よって，μ'_{2r} が存在するための必要十分条件は $-1<2r<k$ のときである．

スチューデント分布 $S(k)$ の平均と分散は，それぞれ $k>1$ と $k>2$ に対して定義され

$$\mathscr{E}(t)=0,\qquad \sigma^2(t)=\frac{k}{k-2}.$$

次の **7.8.4** はスチューデント分布とベータ分布との関係を示している．

7.8.4　t がスチューデント分布 $S(k)$ に従うとき，$x=\dfrac{1}{1+\dfrac{t^2}{k}}$ はベータ分布

$Be\left(\frac{1}{2}k,\frac{1}{2}\right)$ を持つ．

(c)　スネディッカー分布

この分布の数学的な性質とその由来は次に示されている．

7.8.5　u と v が独立な確率変数で，それぞれカイ 2 乗分布 $C(k_1)$，$C(k_2)$ を持つとき，確率変数

$$\mathscr{F}=\frac{u}{k_1}\Big/\frac{v}{k_2}$$

の p.e. は

(7.8.9)
$$dF_{k_1,k_2}(\mathscr{F})=\frac{\Gamma\left(\frac{1}{2}k_1+\frac{1}{2}k_2\right)}{\Gamma\left(\frac{1}{2}k_1\right)\Gamma\left(\frac{1}{2}k_2\right)}(k_1/k_2)^{\frac{1}{2}k_1}\mathscr{F}^{\frac{1}{2}k_1-1}[1+(k_1\mathscr{F}/k_2)]^{-\frac{1}{2}(k_1+k_2)}\,d\mathscr{F}$$

になる．

これが**自由度** k_1，k_2 **のスネディッカー分布**[*]である．簡単に，**スネディッカー分布** $S(k_1,k_2)$[**] と呼ぶこともある．この分布は 10.4 節と 10.6 節において適用されている．

スネディッカー分布は，u と v の p.e.

(7.8.10)
$$\frac{\left(\frac{1}{2}u\right)^{\frac{1}{2}k_1-1}\left(\frac{1}{2}v\right)^{\frac{1}{2}k_2-1}}{4\Gamma\left(\frac{1}{2}k_1\right)\Gamma\left(\frac{1}{2}k_2\right)}e^{-\frac{1}{2}(u+v)}\,du\,dv$$

に変換

(7.8.11)
$$\mathscr{F}=\frac{u}{k_1}\Big/\frac{v}{k_2},\qquad \mathscr{G}=v$$

を行なって，$\mathscr{F}$ の周辺分布をとることによって得られる．

まず，$\mathscr{F}$ と $\mathscr{G}$ の p.e. は

(7.8.12)
$$\frac{\left(\frac{k_1}{k_2}\right)^{\frac{1}{2}k_1}}{2\Gamma\left(\frac{1}{2}k_1\right)\Gamma\left(\frac{1}{2}k_2\right)}\mathscr{F}^{\frac{1}{2}k_1-1}\left(\frac{\mathscr{G}}{2}\right)^{\frac{1}{2}(k_1+k_2)-1}e^{-\frac{1}{2}[1+(k_1\mathscr{F}/k_2)]\mathscr{G}}\,d\mathscr{F}\,d\mathscr{G}.$$

$\mathscr{G}$ で 0 から ∞ まで積分すると，(7.8.9) が $\mathscr{F}$ の p.e. になる．

---

[*] 自由度 k_1，k_2 の F 分布ともいわれている．(訳注)

[**] 自由度 k_1，k_2 の F 分布に対応して，$F(k_1,k_2)$ と書くこともある．(訳注)

スネディッカー分布の r 次のモーメントは

$$\mu_r' = \left(\frac{k_2}{k_1}\right)^r \mathscr{E}(u^r)\mathscr{E}(v^{-r}) = \left(\frac{k_2}{k_1}\right)^r \frac{\Gamma\left(\frac{1}{2}k_1 + r\right)\Gamma\left(\frac{1}{2}k_2 - r\right)}{\Gamma\left(\frac{1}{2}k_1\right)\Gamma\left(\frac{1}{2}k_2\right)}. \tag{7.8.13}$$

これは，$-k_1 < 2r < k_2$ のときに限り存在する．$\mathscr{F}$ の平均と分散は

$$\mu(\mathscr{F}) = \frac{k_2}{k_2 - 2}, \qquad \sigma^2(\mathscr{F}) = \frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)}. \tag{7.8.14}$$

スネディッカー分布もベータ分布と関連している．これを述べよう．

7.8.6　$\mathscr{F}$ がスネディッカー分布 $S(k_1, k_2)$ を持つ確率変数のとき，確率変数 $[1 + k_1\mathscr{F}/k_2]^{-1}$ はベータ分布 $Be\left(\frac{1}{2}k_2, \frac{1}{2}k_1\right)$ に，$\dfrac{k_1\mathscr{F}}{k_2 + k_1\mathscr{F}}$ は $Be\left(\frac{1}{2}k_1, \frac{1}{2}k_2\right)$ に従う．

実際的な立場からすれば，スネディッカー分布は分散分析で用いられる場合，Fisher (1924) がはじめて示した型よりも便利な型になっている．フィッシャーが考えた確率変数 z とスネディッカーの $\mathscr{F}$ とは次の関係になっている：

$$z = \frac{1}{2}\log_e \mathscr{F}. \tag{7.8.15}$$

Fisher (1925 a) は，$P(z > z_\alpha) = \alpha$ なる z_α の値を，$\alpha = 0.01, 0.05$，$k_1 = 1, 2, 3, 4, 5, 6, 8, 12, 24, \infty$，$k_2 = 1, 2, \cdots, 30, 60, \infty$ に対して，表にまとめている．Snedecor (1937) は z_α に対応する $\mathscr{F}_\alpha$ $\left(z_\alpha = \frac{1}{2}\log \mathscr{F}_\alpha\right)$ の値を，$\alpha = 0.01, 0.05$，$k_1 = 1, 2, \cdots, 12, 14, 16, 20, 24, 30, 40, 50, 75, 100, 200, 500, \infty$，$k_2 = 1, 2, \cdots, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 55, 60, 65, 70, 80, 100, 125, 150, 200, 400, 1000, \infty$ について表にしている．詳細は Greenwood と Hartley (1961) を参照せよ．

$\mathscr{F}$ がスネディッカー分布 $S(k_1, k_2)$ を持つとき，$\mathscr{F}_\alpha$ の値は Karl Pearson (1934) の *Tables of the Incomplete Beta Function* から決められることが 7.8.6 からわかる．なぜなら，$P(\mathscr{F} > \mathscr{F}_\alpha) = I_{x_\alpha}\left(\frac{1}{2}k_1, \frac{1}{2}k_2\right)$，$\mathscr{F}_\alpha = \dfrac{k_2 x_\alpha}{k_1(1 - x_\alpha)}$.

問　題

7.1　7.22 を証明せよ．

7.2　x が正規分布 $N(\mu, \sigma^2)$ に従う確率変数のとき

$$\mathscr{E}(|x - \mu|) = \sqrt{(2/\pi)}\,\sigma$$

を示せ．

7.3　x が唯一のメディアン $\underline{x_{0.5}}$ を持つ連続型確率変数ならば，$\mathscr{E}(|x - c|)$ は $c = \underline{x_{0.5}}$ のとき最小になることを示せ．

7.4　x_1 が正規分布 $N(\mu, \sigma^2)$ に従う確率変数で，$x_2|x_1$ が正規分布 $N(x_1, \sigma^2)$ を持つ条件つき確率変数とする．このとき，(x_1, x_2) は 2 次元正規分布 $N(\mu, \mu; \|\sigma_{ij}\|)$ に従うことを示せ．ただし

$$\|\sigma_{ij}\| = \begin{Vmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & 2\sigma^2 \end{Vmatrix}.$$

7.5　$f(x_1, x_2)$ が 2 次元円形正規分布 $N(\{0\}, \|\delta_{ij}\|)$，$i, j = 1, 2$ の p.d.f. のとき，$f(x_1, x_2)$ の正方形：$(-k, -k)$，$(-k, +k)$，$(k, -k)$，(k, k) 上の積分は，円板：$x_1^2 + x_2^2 < 4k^2/\pi$ 上の積分より小さいことを示せ．すなわち

$$\frac{1}{\sqrt{2\pi}}\int_0^k e^{-\frac{1}{2}t^2}\,dt < \frac{1}{2}\sqrt{1 - e^{-2k^2/\pi}}$$

を示せ．これは Williams (1946) による．

7.6　$f(x_1, x_2)$ は 2 次元正規分布 $N(\{0\}, \|\sigma_{ij}\|)$ の p.d.f. とする．ただし

$$\|\sigma_{ij}\| = \begin{Vmatrix} 1 & \rho \\ \rho & 1 \end{Vmatrix}.$$

$F_1(x)$，$F_2(x)$ をそれぞれ x_1，x_2 の周辺分布の c.d.f. とする．確率変数 $F_1(x_1)$，$F_2(x_2)$ の相関係数は $(6/\pi)\sin^{-1}\left(\frac{1}{2}\rho\right)$ であることを示せ．

7.7　(x_1, x_2) は正規分布 $N(\{0\}, \|\sigma_{ij}\|)$ を持つ 2 次元確率変数である．ただし

$$\|\sigma_{ij}\| = \begin{Vmatrix} 1 & \rho \\ \rho & 1 \end{Vmatrix}.$$

このとき，比 $z = x_1/x_2$ の p.d.f. は $\sqrt{1 - \rho^2}/\pi[1 - 2\rho z + z^2]$ になることを示せ．[Fieller (1932)].

7.8　x_1，x_2 がおのおの分布 $R\left(\frac{1}{2}, 1\right)$ を持つ独立な確率変数ならば，$\sqrt{-2\log x_1}\cos 2\pi x_2$，$\sqrt{-2\log x_1}\sin 2\pi x_2$ はおのおの正規分布 $N(0, 1)$ に従い，独立になることを示せ．[Box と Muller (1958)].

7.9　x_1, x_2 がそれぞれガンマ分布 $G(k_1)$，$G(k_2)$ を持つ独立な確率変数ならば，$x_1 + x_2$ と $x_1/(x_1 + x_2)$ は独立で，それぞれガンマ分布 $G(k_1 + k_2)$，ベータ分布 $Be(k_1, k_2)$ に従うことを示せ．

7.10　x_1，x_2 は独立な確率変数で，それぞれガンマ分布 $G(\nu)$，$G\left(\nu + \frac{1}{2}\right)$ に従うとき，$y = 2\sqrt{x_1 x_2}$ はガンマ分布 $G(2\nu)$ を持つことを示せ．

7.11　2 次元確率変数 (x_1, x_2) が x_1x_2 平面で p.d.f.

$$f(x_1, x_2) = \begin{cases} \dfrac{1}{\Gamma(k_1)\Gamma(k_2)} x_1^{k_1-1}(x_2 - x_1)^{k_2-1}e^{-x_2}, & 0 < x_1 < x_2 < \infty \\ 0, & \text{その他} \end{cases}$$

を持つとき，x_1，x_2 の周辺分布がそれぞれガンマ分布 $G(k_1)$，$G(k_1 + k_2)$ になることを

示せ.

7.12　x は, ある正の整数 k に対して, kx がガンマ分布 $G(k)$ を持つような連続型確率変数で, $y|x$ はポアソン分布 $Po(x)$ に従う離散型条件つき確率変数であるとする. y の（非条件つき）分布は 2 項待ち時間分布になることを示せ.

7.13　x_1, x_2 はそれぞれベータ分布 $Be(\nu_1, \nu_2)$, $Be\left(\nu_1 + \frac{1}{2}, \nu_2\right)$ を持つ独立な確率変数である. $\sqrt{x_1 x_2}$ がベータ分布 $Be(2\nu_1, 2\nu_2)$ に従うことを示せ.

7.14　おのおのが矩形分布 $R\left(\frac{1}{2}, 1\right)$ に従う独立な確率変数を $x_1, \cdots, x_k$ とする. このとき, $-\log(x_1 \cdots x_k)$ はガンマ分布 $G(k)$ を持つことを示せ.

7.15　u がカイ 2 乗分布 $C(k)$ に従う確率変数ならば, 特性関数を用いて, $k \longrightarrow \infty$ のとき, $(u-k)/\sqrt{2k}$ は分布 $N(0,1)$ に近づくことを示せ.

7.16　確率変数 x の c.d.f. が $F(x)$ で, その特性関数を $\varphi(t)$ とする. すべての正の整数 n について, $[\varphi(t)]^{\frac{1}{n}}$ がある確率変数の特性関数になるとき, x は**無限分解可能**であるといわれる. すなわち, 各 n に対して, $F(x)$ が, おのおの特性関数 $[\varphi(t)]^{\frac{1}{n}}$ を持つ n 個の独立な確率変数の和の分布になる場合である. ポアソン分布, ガンマ分布, 正規分布に従う確率変数はいずれも無限分解可能になることを示せ. ［無限分解可能な確率変数の取り扱いについては, Gnedenko と Kolmogorov (1954) を参照せよ］.

7.17　(7.8.4) で与えられたスチューデントの t 分布の p.d.f. は, $k \longrightarrow \infty$ のとき, すべての t に対して, 分布 $N(0,1)$ の p.d.f. に収束することを示せ.

7.18　$x_1, \cdots, x_k$ は独立で, すべて分布 $N(\mu, \sigma^2)$ に従い, $c_1, \cdots, c_k$ は和と 2 乗和がともに 1 である（実数）定数である. このとき, $c_1 x_1 + \cdots + c_k x_k$ は分布 $N(\mu, \sigma^2)$ を持つことを示せ. すべてが正で, この条件を満たす c の組は存在しないことを示せ.

7.19　$x_1, \cdots, x_k$ が独立で, おのおの $N(0,1)$ に従う確率変数とする. $y_p = c_{p1}x_1 + \cdots + c_{pk}x_k$, $p = 1, \cdots, s$ で $y_1, \cdots, y_s$, $s \leqslant k$ を定義する. ただし $\sum_{i=1}^{k} c_{pi} c_{qi} = 0$ $(p \neq q)$, 1 $(p = q)$. $y_1, \cdots, y_s$ は分布 $N(0,1)$ を持つ独立な確率変数であることを示せ.

7.20　分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$ を持つ k 次元確率変数 $(x_1, \cdots, x_k)$ の特性関数 (7.4.17) を微分することにより, $\mathscr{E}(x_i) = \mu_i$, $\mathscr{E}(x_i - \mu_i)(x_j - \mu_j) = \sigma_{ij}$ であることを示せ.

7.21　$(x_1, \cdots, x_k)$, $(x_1', \cdots, x_k')$ がそれぞれ $N(\{\mu_i\}, \|\sigma_{ij}\|)$, $N(\{\mu_i'\}, \|\sigma_{ij}'\|)$ を持つ独立な k 次元確率変数であるとき, $(x_1 + x_1', \cdots, x_k + x_k')$ は正規分布 $N(\{\mu_i + \mu_i'\}, \|\sigma_{ij} + \sigma_{ij}'\|)$ に従うことを（特性関数を用いて）示せ. すなわち, k 次元正規分布は平均と共分散行列をこみにしたベクトルに関して再生的である.

7.22　μ をある標準的な金属棒の"真"の長さとする. x_1 はこの棒の最初の写しの長さを表わす確率変数とし, x_2 は x_1 の写しの長さを表わす確率変数とする. すなわち, x_2 は標準的な金属棒の 2 番目の写しの長さを表わしている. 以下, x_k は標準的な棒の k 番目の写しを示す確率変数とする. $(x_1 - \mu), (x_2 - x_1), \cdots, (x_k - x_{k-1})$ はおのおの正規分布 $N(0,1)$ を持ち, 独立であるならば, $(x_1, \cdots, x_k)$ の分布は k 次元正規分布 $N(\{\mu_i\}, \|\sigma_{ij}\|)$ になることを示せ. ただし, $\mu_i = \mu$, $i = 1, \cdots, k$

$$
\|\sigma_{ij}\| = \left\|\begin{matrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 2 & \cdots & 2 & 2 \\ 1 & 2 & 3 & \cdots & 3 & 3 \\ \vdots & \vdots & \vdots & & \vdots & \\ 1 & 2 & 3 & \cdots & k-1 & k-1 \\ 1 & 2 & 3 & \cdots & k-1 & k \end{matrix}\right\|.
$$

7.23　$(x_1, \cdots, x_k)$ は k 次元球形正規分布 $N(\{\mu_i\}, \|\delta_{ij}\sigma^2\|)$ を持つ確率変数である. ただし, $\delta_{ij} = 1$ $(i = j)$, 0 $(i \neq j)$. このとき,

$$r = [(x_1 - \mu_1)^2 + \cdots + (x_k - \mu_k)^2]^{1/2}$$

の平均は

$$\mathscr{E}(r) = \sqrt{2}\,\sigma \frac{\Gamma\left(\frac{1}{2}k + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}k\right)}$$

であることを示せ.

7.24　ある種の個体をランダムに取り出してきたとき, 時間区間 $(t, t + dt)$ で死滅する確率が $\dfrac{t^{k-1}e^{-t}}{\Gamma(k)}\,dt$ であるとする. この個体が死滅するとすぐに別の種類の個体になるとする. 2 番目の個体が死滅するや否や 3 番目の生物になり, 以下, 同様に変化すると仮定する. r 番目の個体が時間 $(t, t + dt)$ で死滅する確率は

$$\frac{t^{rk-1}e^{-t}}{\Gamma(rk)}\,dt$$

になることを示せ.

7.25（続き）　**一般死滅法則を持つ再生過程**　$f(t)dt$ は, 最初の (0) 世代の個体が, 時間 $(t, t + dt)$ の間に次の世代に変わる確率とし, $g_n(\tau)\,d\tau$ は n 番目の世代から, 時間 $(\tau, \tau + d\tau)$ で次の世代に変化する確率とする. どの世代においても, 最初の世代と同じ死滅法則に従うとすれば

$$g_{n+1}(t) = \int_0^t g_n(\tau) f(t - \tau)\,d\tau$$

となることを示せ. したがって, 安定状態になる法則 $g(t)$ が存在するならば, $g(t)$ は積分方程式

$$g(t) = \int_0^t g(\tau) f(t - \tau)\,d\tau$$

を満たさねばならない. [Lotka (1939) と Smith (1958)].

7.26　$x_1, \cdots, x_{k+1}$ は独立で, すべて分布 $N(0,1)$ に従うとする.

$$y_i = \frac{x_i^2}{x_1^2 + \cdots + x_{k+1}^2}, \qquad i = 1, \cdots, k$$

ならば, $(y_1, \cdots, y_k)$ は k 次元ディリクレ分布 $D\left(\frac{1}{2}, \cdots, \frac{1}{2}; \frac{1}{2}\right)$ を持つことを示せ.

7.27　**7.7.1** を証明せよ.

7.28　τ はスネディッカー分布 $S(k_1,k_2)$ を持つ確率変数であるとする．いま $y_{k_2}=k_1\tau$ とすれば，確率変数列 $(y_1,y_2,\cdots)$ は自由度 k_1 のカイ2乗分布に従う確率変数に法則収束することを証明せよ．

7.29　$(x_1,\cdots,x_k)$ は k 次元確率変数で，p.d.f. は

$$f(x)=\begin{cases} g(a_1x_1+\cdots+a_kx_k), & x_1>0,\cdots,x_k>0 \\ 0, & \text{その他} \end{cases}$$

なる型をしているとする．ただし，$a_1,\cdots,a_k$ は正の定数．確率変数 $y=a_1x_1+\cdots+a_kx_k$ の p.d.f. は

$$f(y)=\begin{cases} \dfrac{1}{(a_1\cdots a_k)\Gamma(k)}\,y^{k-1}g(y), & y\geqslant 0 \\ 0, & y<0 \end{cases}$$

で与えられることを示せ．

7.30　$(x_1,\cdots,x_k)$ は $g(y)$ を p.d.f. に持つ k 次元確率変数とする．ただし $y=\sum_{i,j}a_{ij}x_ix_j$ で，$\|a_{ij}\|$ は対称，正定値である．このとき，y の p.d.f. は

$$f(y)=\begin{cases} \dfrac{\pi^{k/2}}{\sqrt{|a_{ij}|}\,\Gamma(k/2)}\,y^{(k/2)-1}g(y), & y\geqslant 0 \\ 0, & y<0 \end{cases}$$

になることを示せ．

7.31　$(x_1,\cdots,x_k)$ は $N(\{\mu_i\};\|\sigma_{ij}\|)$ に従う k 次元確率変数である．条件つき p.d.f. $f(x_1,\cdots,x_s|x_{s+1},\cdots,x_k)$ は s 次元分布 $N(\mu_1^*,\cdots,\mu_s^*;\|\sigma_{pq}^*\|)$ の p.d.f. になることを示せ．ただし μ_p^* は方程式

$$\begin{vmatrix} (\mu_p^*-\mu_p) & \sigma_{p\,s+1} & \cdots & \sigma_{pk} \\ (x_{s+1}-\mu_{s+1}) & \sigma_{s+1\,s+1} & \cdots & \sigma_{s+1\,k} \\ \vdots & \vdots & & \vdots \\ (x_k-\mu_k) & \sigma_{k\,s+1} & \cdots & \sigma_{kk} \end{vmatrix}=0$$

の解である $(p=1,\cdots,s)$．σ_{pq}^* は

$$\sigma_{pq}^*=\frac{\begin{vmatrix} \sigma_{pq} & \sigma_{p\,s+1} & \cdots & \sigma_{pk} \\ \sigma_{s+1\,q} & \sigma_{s+1\,s+1} & \cdots & \sigma_{s+1\,k} \\ \vdots & \vdots & & \vdots \\ \sigma_{kq} & \sigma_{k\,s+1} & \cdots & \sigma_{kk} \end{vmatrix}}{\begin{vmatrix} \sigma_{s+1\,s+1} & \cdots & \sigma_{s+1\,k} \\ \vdots & & \vdots \\ \sigma_{k\,s+1} & \cdots & \sigma_{kk} \end{vmatrix}}$$

で与えられる $(p,q=1,\cdots,s)$．

7.32　ポアソン過程　(7.5.11) について次のように考える．時間 $(0,t+\Delta t)$ の間に C が k 個起る事象を E とすると，E は $(k+1)$ 個の互いに排反な事象 $E_0,E_1,\cdots,$

E_k の和になる．ただし E_i は $(t,t+\Delta t)$ で C が i 回起り，$(0,t)$ で $(k-i)$ 回起る事象とする．したがって

$$P(E)=P(E_0)+P(E_1)+\cdots+P(E_k).$$

時間区間 I で任意回 C が起ることと，$I\cap I'=\varnothing$ なる I' で，C が任意回起ることとは独立であると仮定する．よって，$(t,t+\Delta t)$ で C が $0,1,\cdots,k$ 回起る確率は，それぞれ $1-\lambda\Delta t+o(\Delta t),\ \lambda\Delta t+o(\Delta t),\cdots,(\lambda\Delta t)^k+o((\Delta t)^k)$ である．$P(E)$ を $f_k(t+\Delta t)$ で表わすと

$$P(E_0)=f_k(t)(1-\lambda\Delta t)+o(\Delta t),\qquad P(E_1)=f_{k-1}(t)(\lambda\Delta t)+o(\Delta t)$$

で，$P(E_2),\cdots,P(E_k)$ の次数は $(\Delta t)^2$ またはこれより大きい．ゆえに

$$f_k(t+\Delta t)=f_k(t)(1-\lambda\Delta t)+f_{k-1}(t)\lambda\Delta t+o(\Delta t)$$

より，微分方程式系

$$f_k'(t)=\lambda f_{k-1}(t)-\lambda f_k(t),\quad (k=1,\ f_{-1}=0)\quad k=1,2,\cdots$$

を得る．この微分方程式の解 $f_k(t)$ が (7.5.11) を表わすことを示せ．$\{f_k(t):t>0\}$ は連続パラメータ t を持つ確率過程の例である．この確率過程をポアソン過程という．

7.33　Yule (1924) の出生過程　ある生物の集団は新しい個体を生成することはできるが死滅しないとする．$f_k(t)$ は時刻 t で k 個の個体がある確率とする．$t+\Delta t$ で k 個になる事象 E の確率を考えよう．E は $(k+1)$ 個の排反な事象 $E_0,E_1,\cdots,E_k$ に分割される．ただし E_i は時刻 t では $(k-i)$ 個で，$(t,t+\Delta t)$ 間で i 個生成される事象である．重複しない2つの時間区間での生成は，問題 7.32 と同様に，独立であると仮定する．$f_k(t)$ は微分方程式系

$$f_k'(t)=(k-1)\lambda f_{k-1}(t)-k\lambda f_k(t),\qquad k=1,2,\cdots$$

を満たすことを示せ．

$k=m,\ t=0$ のとき，$f_m(0)=1,\ f_k(0)=0,\ k=m,m+1,\cdots$ ならば，微分方程式の解は

$$f_k(t)=\binom{k-1}{m-1}e^{-m\lambda t}(1-e^{-\lambda t})^{k-m}$$

で与えられることを示せ．

7.34　単純待ち行列　単位時間内に窓口に客が到着する平均回数を λ，同じく（サービスを受けて）窓口を出る客の平均回数を μ とする．$f_k(t)$ を，時刻 t で，k 人の客の待ち行列[*)]ができる確率とする．$f_k(t+\Delta t)$ は時刻 $t+\Delta t$ で k 人の客が並んでいる確率である．E を，時刻 $t+\Delta t$ で，サービスを受ける客が k 人待っている事象とし，$E_{ij(k-i+j)}$ は時刻 t では $k-i+j$ 人の客がいて，$(t,t+\Delta t)$ 間に i 人到着して，j 人（サービスを受けて）帰る事象とすれば，事象 $\{E_{ij(k-i+j)}\}$ は互いに排反で，和が E になる．ただし (i,j) は $k-i+j\geqslant 0$ となる負でない整数 $i,\ j$ の組を動くとする．し

*)　ここでの待ち行列は，サービスを待っている客と受けている客の両方をいう．待っている客のみを待ち行列という場合もある．

たがって

$$f_k(t+\Delta t)=P(E)=\sum_{i,j}P(E_{ij(k-i+j)}).$$

重複しない時間区間では，窓口に来て去るまでは，問題 7.32 と同様に独立とする．ゆえに，無視できない確率を持つ事象は $E_{00(k)}, E_{10(k-1)}, E_{01(k+1)}$ で，この確率は $f_k(t)[1-(\lambda+\mu)\Delta t]+0(\Delta t)$, $f_{k-1}(t)\lambda\Delta t+0(\Delta t)$, $f_{k+1}(t)\mu\Delta t+0(\Delta t)$ である．他の事象の確率は $(\Delta t)^2$ またはこれ以上の次数である．$f_k(t)$ は次の微分方程式系

$$f'_k(t)=[\lambda f_{k-1}(t)+\mu f_{k+1}(t)-(\lambda+\mu)f_k(t)],\qquad k=1,2\cdots$$

を満たすことを示せ．[待ち行列の詳細については，Feller (1957), Kendall (1951, 1953), Morse (1958) を参照せよ．モースの本は文献中心である]．

7.35 (続き) $f'_k(t)=0$ となる安定状態では，問題 7.34 の微分方程式系 (ただし $f_{-1}=0$)

$$\lambda f_{k-1}+\mu f_{k+1}-(\lambda+\mu)f_k=0$$

の解は

$$f_k=\rho^k f_0$$

となることを示せ．ただし $\rho=\lambda/\mu$.

もしこの窓口で長さ n の待ち行列を処理できるならば

$$f_0=\frac{(1-\rho)}{(1-\rho^{n+1})}$$

となることを示せ．安定状態では，待ち行列の長さの平均値 $\mathscr{E}(k)$ は

$$\mathscr{E}(k)=\frac{\rho-(n+1)\rho^{n+1}+n\rho^{n+2}}{(1-\rho)(1-\rho^{n+1})}$$

で与えられることを示せ．また，$\rho=1$ (到着する割合と窓口を出る割合が等しい) に対して，待ち行列の平均の長さは $\frac{1}{2}n$, いくらでも多くの人がサービスを受けられる場合 $(n=\infty)$，この平均値は $\rho/(1-\rho)$ ($\rho<1$ のとき) になることを示せ．さらに，サービスを受けるために l 人の客がいる平均の回数 $\mathscr{E}(l)$ は

$$\mathscr{E}(l)=\sum_{i=1}^{n}(i-1)f_i=\frac{\rho^2-n\rho^{n+1}+(n-1)\rho^{n+2}}{(1-\rho)(1-\rho^n)}$$

で与えられることを示せ．

第8章 標本論

8.1 確率標本の定義

x が R_1 上の c.d.f. $F(x)$ を持つ1次元確率変数であるとき，c.d.f. $F(x)$ を持つ母集団からとった大きさ n の確率標本は，標本空間 $R_n=R_1^{(1)}\times\cdots\times R_1^{(n)}$ における c.d.f.

(8.1.1)
$$\prod_{\xi=1}^{n}F(x_\xi)$$

を持つ n 次元確率変数として定義される．ただし，$R_1^{(\xi)}$ は x_ξ, $\xi=1,\cdots,n$ の1次元標本空間 (実数空間) である．ここで，特に標本の要素または成分 $x_1,\cdots,x_n$ は互いに独立で，すべて同じ c.d.f. を持つことに注意しなければならない．数理統計学では，上で定義したような確率標本の概念は Fisher (1915) が最初に紹介したが，実際には R_n に "標本空間" という用語を用いなかった．"標本空間" という言葉は，1940年までこの意味で用いられている．確率標本の "確率" という言葉は，主に表現の効果をねらっているに過ぎず，一般に，標本と呼ばれている．その後，第1章で用いられているように広義に解釈されている．この意味から，以後，単に標本と呼び，標本 $(x_1,\cdots,x_n)$ を O_n と略記する．$(x_1,\cdots,x_n)$ (または O_n) を簡単に，$F(x)$ からの大きさ n の標本という．$F(x)$ からの大きさ n の標本は有限確率過程の簡単な例になっているが，ときどき，単純無作為抽出とも呼ばれている．

x が p.f. $p(x)$ を持つ離散型確率変数のとき，その標本は R_n において p.f.

(8.1.2)
$$\prod_{\xi=1}^{n}p(x_\xi)$$

を持つ．x が連続型のときは，R_n において p.e.

(8.1.3)
$$f(x_1)\cdots f(x_n)\,dx_1\cdots dx_n$$

を持つ．連続型の場合と同様に，離散型でも $(x_1, \cdots, x_n)$ を $p(x)$ からの標本という．

注意　x_1 は母集団から最初に取り出した x の値を表わす確率変数，x_2 は2度目に取り出した x の値を示す確率変数，…，と考えると良い．たとえば，"正しい" サイコロを n 回続けて投げる場合，x_ξ は ξ 回目 $(\xi = 1, \cdots, n)$ に投げられたサイコロの目の数を表わす確率変数と考えられる．このとき，標本 O_n は，おのおの

$$p(x) = \frac{1}{6}, \qquad x = 1, \cdots, 6$$

なる p.f. を持つ独立な確率変数 $x_1, \cdots, x_n$ からなっている．標本 O_n の p.f.

$$p(x_1)\cdots p(x_n) = \left(\frac{1}{6}\right)^n$$

は R_n の 6^n 個の質点で定義されている．6^n 個の点は，サイコロを n 回投げて得られる目の可能な列が 6^n 個あることに対応している．

標本論では，標本を構成している n 個の確率変数の1つの関数，あるいは複数個の関数が問題になる．たとえば，標本和 $z = \sum_{\xi=1}^{n} x_\xi$，標本平均 $\bar{x} = \frac{1}{n}\sum_{\xi=1}^{n} x_\xi$，標本分散 $s^2 = \frac{1}{n-1}\sum_{\xi=1}^{n}(x_\xi - \bar{x})^2$，標本最小値 $\min(x_1, \cdots, x_n)$，標本最大値 $\max(x_1, \cdots, x_n)$ などがある．ここでは一般に確率変数 $g(x_1, \cdots, x_n)$ の分布関数を定めるのが問題である．$g(x_1, \cdots, x_n)$ を統計量といい，この c.d.f. $H(y)$ は (2.8.12) の特別な場合である．すなわち

(8.1.4)　$$H(y) = P(g(x_1, \cdots, x_n) \leqq y) = \int_{g^{-1}(y)} dF(x_1)\cdots dF(x_n).$$

ただし，$g^{-1}(y)$ は $g(x_1, \cdots, x_n) \leqq y$ なる R_n の部分集合である．関数 $H(y)$ は $g(x_1, \cdots, x_n)$ の c.d.f. である．

$g_i(x_1, \cdots, x_n)$，$i = 1, \cdots, s \leqq n$ が s 個の（関数として独立な）統計量のときも，同様に $(y_1, \cdots, y_s)$ なる統計量の分布が問題になる．$g_i(x_1, \cdots, x_n)$ の c.d.f. は

(8.1.5)　$$\begin{aligned} H(y_1, \cdots, y_s) &= P(g_i(x_1, \cdots, x_n) \leqq y_i,\ i = 1, \cdots, s) \\ &= \int_{g^{-1}(y_1, \cdots, y_s)} dF(x_1)\cdots dF(x_n) \end{aligned}$$

で定義される．ただし $g^{-1}(y_1, \cdots, y_s)$ は

$$g_i(x_1, \cdots, x_n) \leqq y_i, \qquad i = 1, \cdots, s$$

なる R_n の部分集合である．

確率変数 x が R_k における c.d.f. $F(x_1, \cdots, x_k)$ を持つ k 次元確率変数 $(x_1, \cdots, x_k)$ のときは，標本 O_n は nk 次元確率変数 $(x_{1\xi}, \cdots, x_{k\xi}; \xi = 1, \cdots, n)$ で，$R_{nk} = R_k^{(1)} \times \cdots \times R_k^{(n)}$ において c.d.f.

(8.1.6)　$$\prod_{\xi=1}^{n} F(x_{1\xi}, \cdots, x_{k\xi})$$

を持つ．ここで $R_k^{(\xi)}$ は $(x_{1\xi}, \cdots, x_{k\xi})$，$\xi = 1, \cdots, n$ の標本空間である．ここでも，nk 個の確率変数 $x_{i\xi}$，$i = 1, \cdots, k$；$\xi = 1, \cdots, n$ の1つの関数，あるいは複数個の関数の分布を定める問題がある．たとえば，標本和 $z_i = \sum_{\xi=1}^{n} x_{i\xi}$，標本平均 $\bar{x}_i = \frac{1}{n}\sum_{\xi=1}^{n} x_{i\xi}$，標本共分散 $s_{ij} = \frac{1}{n-1}\sum_{\xi=1}^{n}(x_{i\xi} - \bar{x}_i)(x_{j\xi} - \bar{x}_j)$ などの統計量が考えられる．

標本論では2つあるいはそれ以上の標本の関数を取り扱わねばならないことがある．たとえば，$O_{n_1} : (x_{11}, \cdots, x_{1n_1})$，$O_{n_2} : (x_{21}, \cdots, x_{2n_2})$ がそれぞれ c.d.f. $F_1(x)$，$F_2(x)$ を持つ大きさ n_1，n_2 の標本であるとき，$(x_{11}, \cdots, x_{1n_1}, x_{21}, \cdots, x_{2n_2})$ は $R_{n_1+n_2}$ において c.d.f.

(8.1.7)　$$\prod_{\xi_1=1}^{n_1} F_1(x_{1\xi_1}) \cdot \prod_{\xi_2=1}^{n_2} F_2(x_{2\xi_2})$$

を持つ $(n_1 + n_2)$ 次元確率変数と考えられる．$(n_1 + n_2)$ 次元確率変数の各成分の1つの関数，あるいは複数個の関数の分布を求めたり，少なくともこの分布の性質を明らかにする必要があるだろう．さらに，3つないしそれ以上の標本についても同様である．

数理統計学における標本分布では，特に，平均，2乗和，尤度比，共分散などのように比較的簡単な統計量がしばしば問題になる．このような標本分布の p.f. や p.d.f. は，母集団分布が次で見るように特殊な分布に限られるときは容易に表わせる．しかし，3.3, 3.4, 3.5 の結果を適用すれば，一般的な母集団に対しても，この統計量の平均，分散，共分散，低次のモーメントなどを求めることができる．次節で，この種の重要な結果を考えよう．

8.2　標本平均，標本分散および標本（対称）関数の期待値*) と分散

(a)　標本平均の期待値と分散

$(x_1, \cdots, x_n)$ は平均 μ，分散 σ^2 を持つ母集団からの標本とする．この標本平均 $\bar{x}$ の期待値と分散について考えよう．定義より

(8.2.1)　$$\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n).$$

*) 本来ならば平均（値）であるが，ここでは標本平均の平均（値）といういい方を避けるために，標本平均の期待値というように，"期待値" という用語を用いた．（訳注）

この両辺の期待値をとると

(8.2.2) $$\mathscr{E}(\bar{x})=\frac{1}{n}[\mathscr{E}(x_1)+\cdots+\mathscr{E}(x_n)].$$

しかも，$x_1,\cdots,x_n$ は同一の c.d.f. を持つ確率変数だから

$$\mathscr{E}(x_1)=\cdots=\mathscr{E}(x_n)=\mu.$$

これを (8.2.2) に代入すると

(8.2.3) $$\mathscr{E}(\bar{x})=\mu$$

が得られる．

ここでは，統計的推定に関する基礎概念や原理には直接触れないで，(8.2.3) が成り立つ場合，$\bar{x}$ は μ の**不偏推定量**である，といっておく．推定量については，第10，11，12章で詳しく議論する．

さて，$\bar{x}$ の分散は

$$\sigma^2(\bar{x})=\mathscr{E}[\bar{x}-\mathscr{E}(\bar{x})]^2.$$

一方，
$$\bar{x}-\mathscr{E}(\bar{x})=\frac{1}{n}\sum_{\xi}(x_\xi-\mu).$$

よって，
$$\sigma^2(\bar{x})=\mathscr{E}\Big[\frac{1}{n}\sum_{\xi}(x_\xi-\mu)\Big]^2$$
$$=\frac{1}{n^2}\sum_{\xi}\mathscr{E}(x_\xi-\mu)^2+\frac{1}{n^2}\sum_{\xi\neq\eta}\mathscr{E}[(x_\xi-\mu)(x_\eta-\mu)].$$

$x_1,\cdots,x_n$ は独立で同じ分布に従うから

(8.2.4) $$\begin{aligned}&\mathscr{E}(x_\xi-\mu)^2=\sigma^2, && \xi=1,\cdots,n\\ &\mathscr{E}[(x_\xi-\mu)(x_\eta-\mu)]=0, && \xi\neq\eta.\end{aligned}$$

ゆえに，

(8.2.5) $$\sigma^2(\bar{x})=\frac{\sigma^2}{n}.$$

(b) 標本分散の期待値と分散

さて，

(8.2.6) $$s^2=\frac{1}{n-1}\sum_{\xi=1}^{n}(x_\xi-\bar{x})^2$$

で定義される標本分散は

(8.2.6 a) $$s^2=\frac{1}{n-1}\sum_{\xi}\Big[(x_\xi-\mu)-\frac{1}{n}\sum_{\eta}(x_\eta-\mu)\Big]^2$$
$$=\frac{1}{n}\sum_{\xi}(x_\xi-\mu)^2-\frac{1}{n(n-1)}\sum_{\xi\neq\eta}(x_\xi-\mu)(x_\eta-\mu)$$

となる．両辺の期待値をとって，(8.2.4) を用いると

(8.2.7) $$\mathscr{E}(s^2)=\sigma^2.$$

ここで，(8.2.6) で，n のかわりに $(n-1)$ で割った理由が，$\mathscr{E}(s^2)$ を σ^2 に等しくするためであることに気づくであろう．すなわち，s^2 は σ^2 の不偏推定量である．

同様に $(s^2)^2$ の期待値は簡単な式の変形より

(8.2.8) $$\mathscr{E}[(s^2)^2]=\frac{\mu_4}{n}+\frac{(n-1)^2+2}{n(n-1)}\sigma^4$$

となる．ただし μ_4 は母集団分布の4次の中心モーメントである．(8.2.7)，(8.2.8) より

$$\sigma^2(s^2)=\mathscr{E}[(s^2)^2]-[\mathscr{E}(s^2)]^2$$

だから，s^2 の分散は

(8.2.9) $$\sigma^2(s^2)=\frac{1}{n}\Big(\mu_4-\frac{n-3}{n-1}\sigma^4\Big)$$

になる．以上まとめると

8.2.1　$(x_1,\cdots,x_n)$ を平均 μ，分散 σ^2 を持つ母集団からの標本とするとき，標本平均の期待値と分散は

$$\mathscr{E}(\bar{x})=\mu,\qquad \sigma^2(\bar{x})=\frac{\sigma^2}{n}.$$

さらに，母集団分布の4次のモーメントが有限ならば，標本分散 s^2 の期待値と分散は

$$\mathscr{E}(s^2)=\sigma^2,\qquad \sigma^2(s^2)=\frac{1}{n}\Big(\mu_4-\frac{n-3}{n-1}\sigma^4\Big)$$

である．

高次の標本モーメントの期待値と分散も同様にして考えられるが，実際に計算を行なうのは大変面倒である．詳しく知りたい読者は，Kendall (1943) を参照せよ．

(c) フィッシャーの k 統計量

ある特殊な目的のためには，分布のモーメントよりも半不変係数を取り扱う方が便利な

場合がある．この節では，母集団からの標本の関数が半不変係数の不偏推定量になっている場合を考えよう．

Fisher (1928 a) は次の (i), (ii) を満たす $x_1, \cdots, x_n$ に関する r 次の斉次多項式 $k_r(x_1, \cdots, x_n)$，$r = 1, 2, \cdots$ を考えた．(i) $k_r(x_1, \cdots, x_n)$ は $x_1, \cdots, x_n$ に関して対称である．(ii) $\mathscr{E}(k_r) = \kappa_r$，ただし κ_r は (5.1.11) で定義されたように，標本が取り出された分布の r 次の半不変係数（累積率）である．このような関数を k 統計量という．最初の3つの k 統計量に関しては，多項式の係数が定まることが容易に示される．さらに，k 統計量の構成，その標本論および k 統計量に関する諸々の問題については Craig (1928), Fisher (1928 a), Cornish と Fisher (1937), Dwyer (1938), Kendall (1943) によって詳しく議論されている．

最初の3つの k 統計量は次の型になる．

(8.2.10)
$$\begin{aligned} k_1 &= a_{11}\sum_{\xi} x_\xi \\ k_2 &= a_{21}\sum_{\xi} x_\xi^2 + a_{22}\sum_{\xi \neq \eta} x_\xi x_\eta \\ k_3 &= a_{31}\sum_{\xi} x_\xi^3 + a_{32}\sum_{\xi \neq \eta} x_\xi^2 x_\eta + a_{33}\sum_{\xi \neq \eta \neq \zeta} x_\xi x_\eta x_\zeta. \end{aligned}$$

この統計量の期待値が母集団分布の半不変係数 κ_1, κ_2, κ_3，すなわち

(8.2.11)
$$\begin{aligned} \kappa_1 &= \mu_1' = \mu \\ \kappa_2 &= \mu_2' - (\mu_1')^2 = \sigma^2 \\ \kappa_3 &= \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3 \end{aligned}$$

に等しいことから，多項式の係数が定まる．ただし，μ_1', μ_2', μ_3' は母集団分布の最初の3つのモーメントである．(8.2.1) を (8.2.10) に代入すると，3つの k 統計量は

(8.2.12)
$$\begin{aligned} k_1 &= \frac{1}{n} z_1 = \bar{x} \\ k_2 &= \frac{1}{n(n-1)}(n z_2 - z_1^2) = s^2 \\ k_3 &= \frac{1}{n(n-1)(n-2)}(n^2 z_3 - 3 n z_1 z_2 + 2 z_1^3) \end{aligned}$$

になる．ここに $z_i = \sum_{\xi} x_\xi^i$, $i = 1, 2, 3$.

最初の2つの k 統計量はちょうど，$\bar{x}$, s^2 に等しくなり，その期待値は μ, σ^2 であるが，これはともに最初の2つの半不変係数 κ_1, κ_2 に等しい．

(d) 標本の対称関数の期待値と分散

これまで考察してきた標本平均 $\bar{x}$，標本分散 s^2，標本モーメント，標本 k 統計量はいずれも標本 $(x_1, \cdots, x_n)$ の各要素に関して対称になっていた．ここでは，このような標本の対称関数のうちでより一般的な族を考えよう．

$(x_1, \cdots, x_r)$ を c.d.f. $F(x)$ からの大きさ r の標本とし，$g(x_1, \cdots, x_r)$ は $(x_1, \cdots, x_r)$ の関数で，

(8.2.13)
$$\mathscr{E}(g^i(x_1, \cdots, x_r)) = \begin{cases} \theta_1, & i = 1 \\ \theta_2, & i = 2 \end{cases}$$

を満たすものとする．ただし θ_1, θ_2 はともに有限である．大きさ $n(\geqq r)$ の標本 $(x_1, \cdots, x_n)$ について

(8.2.14)
$$g_0(x_{\eta_1}, \cdots, x_{\eta_r}) = \frac{1}{r!}\sum\nolimits_p g(x_{\xi_1}, \cdots, x_{\xi_r})$$

とする．ただし，$\eta_1, \cdots, \eta_r$ は整数 $1, \cdots, n$ から $\eta_1 < \cdots < \eta_r$ となるように r 個を選んだもので，$\sum_p$ は $r!$ 個の $(\eta_1, \cdots, \eta_r)$ の置換 $(\xi_1, \cdots, \xi_r)$ のすべてにわたる和を表わす．このとき $g_0(x_{\eta_1}, \cdots, x_{\eta_r})$ は $(x_{\eta_1}, \cdots, x_{\eta_r})$ の対称関数になっている．

(8.2.15)
$$Q_{[r]}(x_1, \cdots, x_n) = \binom{n}{r}^{-1}\sum\nolimits_c g_0(x_{\eta_1}, \cdots, x_{\eta_r})$$

と定義すると，$Q_{[r]}(x_1, \cdots, x_n)$ も標本 $(x_1, \cdots, x_n)$ の対称関数である．ただし $\sum_c$ は $(1, \cdots, n)$ から $(\eta_1, \cdots, \eta_r)$ の $\binom{n}{r}$ 個の選び方にわたる和を示す．さて，標本 $(x_1, \cdots, x_n)$ のすべての置換 $(x_{\xi_1}, \cdots, x_{\xi_r})$ に対して，$\mathscr{E}(g(x_{\xi_1}, \cdots, x_{\xi_r}))$ は有限で，すべて θ_1 に等しい．さらに

(8.2.16)
$$\mathscr{E}(g_0(x_{\eta_1}, \cdots, x_{\eta_r})) = \frac{1}{r!}\sum\nolimits_p \mathscr{E}(g(x_{\xi_1}, \cdots, x_{\xi_r})) = \theta_1$$

(8.2.17)
$$\mathscr{E}(Q_{[r]}(x_1, \cdots, x_n)) = \binom{n}{r}^{-1}\sum\nolimits_c \mathscr{E}(g_0(x_{\eta_1}, \cdots, x_{\eta_r})) = \theta_1$$

は明らかである．

さて，$Q_{[r]}(x_1, \cdots, x_n)$ の分散を求めよう．

(8.2.18)
$$\sigma^2(Q) = \mathscr{E}(Q_{[r]}^2(x_1, \cdots, x_n)) - \theta_1^2$$

で，

(8.2.19)
$$\mathscr{E}(Q_{[r]}^2(x_1, \cdots, x_n)) = \binom{n}{r}^{-2}\sum\nolimits_c' \mathscr{E}(g_0(x_{\eta_1}, \cdots, x_{\eta_r})g_0(x_{\zeta_1}, \cdots, x_{\zeta_r}))$$

が成り立つ．$\eta_1, \cdots, \eta_r$ と $\zeta_1, \cdots, \zeta_r$ は $(1, \cdots, n)$ からともに，$\eta_1 < \cdots < \eta_r$, $\zeta_1 < \cdots < \zeta_r$

となるように，r 個を選びだしたもので，$\sum_c'$ はこのようなすべての選び方にわたる和を示している．いま，$\eta_1,\cdots,\eta_r$ と $\zeta_1,\cdots,\zeta_r$ が共通な整数を j 個持っているならば，このような1対の組 $((\eta_1,\cdots,\eta_r),(\zeta_1,\cdots,\zeta_r))$ に対する $\mathscr{E}(g_0(x_{\eta_1},\cdots,x_{\eta_r})g_0(x_{\zeta_1},\cdots,x_{\zeta_r}))$ の値は常にある一定な値 φ_j に等しい．(8.2.13) の θ_1, θ_2 がともに有限だから，シュワルツの不等式より，φ_j, $j=0,1,\cdots,r$ はいずれも有限である．また，このような組合せは $\binom{n}{r}\binom{r}{j}\binom{n-r}{r-j}$ 個ある．j は $0,1,\cdots,r$ のどれかの値をとる．したがって

(8.2.20)
$$\mathscr{E}(Q_{[r]}^2(x_1,\cdots,x_n))=\binom{n}{r}^{-1}\sum_{j=0}^{r}\binom{r}{j}\binom{n-r}{r-j}\varphi_j.$$

ゆえに，

(8.2.21)
$$\sigma^2(Q_{[r]})=\binom{n}{r}^{-1}\sum_{j=0}^{r}\binom{r}{j}\binom{n-r}{r-j}\varphi_j-\theta_1^2.$$

いま，$\theta_1=0$ になるように $g(x_{\xi_1},\cdots,x_{\xi_r})$ を選べば，$\varphi_0=\theta_1^2$ だから

(8.2.22)
$$\sigma^2(Q_{[r]})=\binom{n}{r}^{-1}\sum_{j=1}^{r}\binom{r}{j}\binom{n-r}{r-j}\varphi_j$$

として，$\sigma^2(Q_{[r]})$ が求められる．これは Hoeffding (1948 *a*) による．次にこの結果をまとめておく．

8.2.2　$(x_1,\cdots,x_n)$ は任意の分布からの標本とし，$g(x_1,\cdots,x_r)$ は1次，2次のモーメントとして，ともに有限な θ_1,θ_2 を持つ関数とする．同じ分布からの標本 $(x_1,\cdots,x_r)$, $n\geqq r$ に対して，(8.2.15) で $Q_{[r]}(x_1,\cdots,x_n)$ を定義すると，$Q_{[r]}(x_1,\cdots,x_n)$ は θ_1 の不偏推定量であり，この分散は (8.2.21) で与えられる．

θ_1,θ_2 が有限で，$\theta_2>\theta_1^2$ ならば，$\sqrt{n}(Q_{[r]}-\theta_1)/\sigma(Q_{[r]})$ は $n\longrightarrow\infty$ のとき $N(0,1)$ に近づくことがホエフディングによって示されている．さらに，ホエフディングは **8.2.2** を $Q_{[r]}(x_1,\cdots,x_n)$ がベクトルの場合に拡張している．

8.3　標本和と標本平均

(a)　繰り返し法

まず，標本和の関数および標本平均の関数の分布を求めるために，繰り返し法を考えよ

う．特殊な場合にはこの方法によって，直接アプローチできる．

大きさ2の標本和の分布を考える．z の c.d.f. を $G_2(z)$ とすると

(8.3.1)
$$G_2(z)=P(x_1+x_2\leqq z)=\int_E d(F(x_1)F(x_2)).$$

ただし E は $x_1+x_2\leqq z$ なる R_2 における事象を表わす．x_1, x_2 は独立で，

$$g(x_1,x_2)=\begin{cases}1, & (x_1,x_2)\in E \text{ のとき}\\ 0, & (x_1,x_2)\in\overline{E}\text{ のとき}\end{cases}$$

として，**3.7.2** を適用すると

(8.3.2)
$$\int_E d(F(x_1)F(x_2))=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{z-x_1}dF(x_2)\right]dF(x_1)=\int_{-\infty}^{\infty}F(z-x_1)\,dF(x_1).$$

ゆえに，大きさ2の標本に対しては次式が成り立つ．

(8.3.3)
$$G_2(z)=\int_{-\infty}^{\infty}F(z-x_1)\,dF(x_1).$$

この方法を，大きさ n の標本にまで拡張し，z の c.d.f. を $G_n(z)$ とすれば，$G_n(z)$ は次のように多重積分の形に表わされる．

(8.3.4)
$$G_n(z)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{z-x_1-\cdots-x_{n-3}}\int_{-\infty}^{z-x_1-\cdots-x_{n-2}}F(z-x_1-\cdots-x_{n-1})\cdot dF(x_{n-1})\,dF(x_{n-2})\cdots dF(x_1).$$

標本平均 $\bar{x}$ の c.d.f. $H_n(\bar{x})$ は，関係式 $H_n(\bar{x})\equiv G_n(n\bar{x})$ によって与えられる．したがって，次が成立する．

8.3.1　$(x_1,\cdots,x_n)$ を c.d.f. $F(x)$ からの大きさ n の標本とすれば，標本和 z の c.d.f. $G_n(z)$ は (8.3.4) で与えられる．標本平均 $\bar{x}$ の c.d.f. $H_n(\bar{x})$ は $H_n(\bar{x})\equiv G_n(n\bar{x})$ で与えられる．

c.d.f. $G_n(z)$ は c.d.f. $F(x_1),\cdots,F(x_n)$ のたたみこみと呼ばれている．任意の分布 $F_1(x_1),\cdots,F_n(x_n)$ を持つ n 個の独立な確率変数のたたみこみも同様に計算される．この場合，たたみこみは n 個の独立な確率変数の和の分布になる．

p.f. $p(x)$ なる離散型分布を持つ母集団からの標本のとき，z の p.f. $p_n(z)$ は $p_1(x)\equiv p(x)$ なるとき，次式を満たすことが確かめられる．

(8.3.4 *a*)
$$p_n(z)=\sum_x p(z-x)p_{n-1}(x).$$

p.d.f. $f(x)$ なる連続分布を持つ母集団からの標本の場合，z の p.d.f. $f_n(z)$ も同

様に

(8.3.4 b) $$f_n(z)=\int_{-\infty}^{\infty} f(z-x)f_{n-1}(x)\,dx$$

で与えられる．ただし $f_1(x)\equiv f(x)$.

例題 (8.3.4 *b*) の特別な場合として，$f(x)$ が矩形分布 $R\left(\frac{1}{2},1\right)$ の p.d.f. のとき

$$f_1(z)=f(z)=\begin{cases}0, & z\leqslant 0\\ 1, & 0<z\leqslant 1\\ 0, & z>1\end{cases}$$

$$f_2(z)=\begin{cases}0, & z\leqslant 0\\ z, & 0<z\leqslant 1\\ z-2(z-1), & 1<z\leqslant 2\\ 0, & z>2\end{cases}$$

となり，数学的帰納法により

$$f_n(z)=\frac{1}{(n-1)!}\cdot\left[z^{n-1}-\binom{n}{1}(z-1)^{n-1}+\binom{n}{2}(z-2)^{n-1}-\cdots+(-1)^k\binom{n}{k}(z-k)^{n-1}\right],$$
$$k<z\leqslant k+1,\ k=0,1,\cdots,n-1 \text{ のとき}$$
$$f_n(z)=0,\qquad z\leqslant 0,\ z>n \text{ のとき}$$

が確かめられる．$\bar{x}$ の p.f. は

$$\begin{cases}nf_n(n\bar{x}), & \frac{k}{n}<\bar{x}\leqslant\frac{k+1}{n},\ k=0,1,\cdots,n-1\\ 0, & \text{その他.}\end{cases}$$

この結果は Laplace (1814) による．

(b) 特性関数の応用

多くの場合，特性関数は標本和の関数あるいは標本平均の関数の分布を求めるとき，大変有効で，しかも，簡単に適用できる．

5.3.2 を用いれば，次の **8.3.2** は **5.3.2** の系として，すぐ成り立つことがわかる．

8.3.2 $(x_1,\cdots,x_n)$ を c.d.f. $F(x)$ からの標本，$\varphi(t)$ を $F(x)$ の特性関数とすれば，標本和 z の特性関数は

$$[\varphi(t)]^n.$$

標本平均 $\bar{x}$ の特性関数は

$$\left[\varphi\left(\frac{t}{n}\right)\right]^n$$

になる．

5.1.2 を適用すれば，z および $\bar{x}$ の c.d.f.（連続型の場合は p.d.f.）は，少なくとも理論的にはその特性関数より定められる．しかし，実際には，特殊な場合をのぞけば，(5.1.14)，(5.1.15) の積分の評価は煩雑である．Irwin (1930) は種々の分布からの標本平均の分布を求めるのに，幅広くこの方法を展開している．しかし，標本論上重要で特殊なものについては，多くの場合，c.d.f. の再生性に関する **5.3.3** を用いて，z や $\bar{x}$ の関数の分布を求めることができる．なぜなら，$(x_1,\cdots,x_n)$ が c.d.f. $F(x;\theta)$ からの標本で，特性関数 $\varphi(t;\theta)$ を持つとすれば，**8.3.2** より，標本和 z の特性関数は $[\varphi(t;\theta)]^n$ になる．しかも，$\varphi(t;\theta)$ が (5.3.7) のように θ に関して再生性を持てば，

(8.3.5) $$[\varphi(t;\theta)]^n=\varphi(t;n\theta).$$

よって，z は c.d.f. $F(z;n\theta)$ を持つことになる．$z=n\bar{x}$ と置くと，標本平均 $\bar{x}$ の c.d.f. は $F(n\bar{x};n\theta)$ となる．まとめると

8.3.3 $(x_1,\cdots,x_n)$ が c.d.f. $F(x;\theta)$ からの標本で，$F(x;\theta)$ の特性関数が $\varphi(t;\theta)$ のとき，$[\varphi(t;\theta)]^n=\varphi(t;n\theta)$ ならば，標本和 z，標本平均 $\bar{x}$ の分布の c.d.f. はそれぞれ $F(z;n\theta)$，$F(n\bar{x};n\theta)$ になる．

8.3.3 は k 変量の分布を持つ母集団からの標本に対しても拡張できる．このとき，z はベクトル $(z_1,\cdots,z_k)$ で，$z_i=\sum_{\xi=1}^{n}x_{i\xi}$．パラメータ θ もいくつかの成分からなるベクトルであってもよい．

次の **8.3.3** の系によって，特定の分布に関する標本和の理論について情報が得られる．これは後章で重要となる．

8.3.3 a $(x_1,\cdots,x_n)$ が2項分布 $Bi(m,p)$ からの標本のとき，和 z の標本分布も2項分布 $Bi(mn,p)$ になる．

なぜなら，2項分布 $Bi(m,p)$ の特性関数は $(q+e^{it}p)^m$，よって，z の特性関数は $(q+e^{it}p)^{mn}$ となるから，z の標本分布は $Bi(mn,p)$ になる．$m=1$ のとき，p.f.

(8.3.6) $$p(x)=\begin{cases}q, & x=0\\ p, & x=1\end{cases}$$

を持つ2項母集団からの標本になることに注意せよ．p.f. (6.2.2) を持つ2項分布は，本質的には 0 と 1 をそれぞれ確率 q, p でとる x_ξ の標本和 $z=\sum_{\xi=1}^{n} x_\xi$ の分布である．

8.3.3 b　$(x_{1\xi},\cdots,x_{k\xi}\,;\,\xi=1,\cdots,n)$ が多項分布 $M(m\,;\,p_1,\cdots,p_k)$ からの標本のとき，標本和 $(z_1,\cdots,z_k)$ の分布は多項分布 $M(mn\,;\,p_1,\cdots,p_k)$ になる．

これは **8.3.3 a** と同様に確かめられる．$m=1$ のとき，多項分布 $M(n\,;\,p_1,\cdots,p_k)$ は (k 変量) 多項分布 $M(1\,;\,p_1,\cdots,p_k)$ からの標本和の分布である．ただしこの多項分布は

(8.3.7)　$$p(x_1,\cdots,x_k)=p_1^{x_1}\cdots p_k^{x_k}(1-p_1-\cdots-p_k)^{1-x_1-\cdots-x_k}$$

なる p.f. を持つ．ここで各 x_i は 0 か 1 で，$x_1+\cdots+x_k\leqq 1$ なる確率変数である．

8.3.3 c　$(x_1,\cdots,x_n)$ はポアソン分布 $Po(\mu)$ からの標本である．このとき z の標本分布はポアソン分布 $Po(n\mu)$ になる．

なぜなら，ポアソン分布 $Po(\mu)$ の特性関数は $e^{-\mu(1-e^{it})}$ だから，z の特性関数は $e^{-n\mu(1-e^{it})}$．これはポアソン分布 $Po(n\mu)$ の特性関数にほかならない．ゆえに，z はポアソン分布 $Po(n\mu)$ を持つ．

8.3.3 d　$(x_1,\cdots,x_n)$ を正規分布 $N(\mu,\sigma^2)$ からの標本とする．z の標本分布は $N(n\mu, n\sigma^2)$ である．$\bar{x}$ の標本分布は $N(\mu,\sigma^2/n)$ である．

なぜなら 7.2 節より，正規分布 $N(\mu,\sigma^2)$ の特性関数は $e^{i\mu t-\frac{1}{2}\sigma^2t^2}$．一方，$z$ は $e^{i(n\mu)t-\frac{1}{2}(n\sigma^2)t^2}$ を特性関数に持つ．これは $N(n\mu,n\sigma^2)$ の特性関数，したがって，z の標本分布の特性関数である．$\bar{x}$ の特性関数としては，z の特性関数における t を t/n に置き換えることにより，$e^{i\mu t-\frac{1}{2}(\sigma^2/n)t^2}$ を得る．これは $N(\mu,\sigma^2/n)$ の特性関数，したがって，$\bar{x}$ の特性関数である．

k 次元正規分布の場合は次のように述べられる．

8.3.3 e　$(x_{1\xi},\cdots,x_{k\xi}\,;\,\xi=1,\cdots,n)$ を k 変数正規分布 $N(\{\mu_i\},\|\sigma_{ij}\|)$ からの標本とする．和 $(z_1,\cdots,z_k)$ の標本分布は $N(\{n\mu_i\},\|n\sigma_{ij}\|)$ で，標本平均 $(\bar{x}_1,\cdots,\bar{x}_k)$ の分布は $N(\{\mu_i\},\|\sigma_{ij}/n\|)$ になる．

これも **8.3.3 d** と同様に確かめられる．各自試みよ．

最後に **8.3.3** の系として，ガンマ分布について述べておく．これも読者自身で確かめよ．

8.3.3 f　$(x_1,\cdots,x_k)$ がガンマ分布 $G(\mu)$ からの標本のとき，z の標本分布はガンマ分布 $G(n\mu)$ である．

取り出された標本 $(x_1,\cdots,x_n)$ の母集団の分布関数が再生的でないとき，一般に，特性関数は標本和や標本平均の分布を求めるためには，あまり役立たない．

標本論では，しばしば，正規母集団からの標本平均の線形関数を取り扱うことがある．次にこのような関数の分布について有効な結果を述べておこう．

8.3.4　$(x_{11},\cdots,x_{1n_1}),\cdots,(x_{k1},\cdots,x_{kn_k})$ はそれぞれ正規分布 $N(\mu_1,\sigma_1^2),\cdots,N(\mu_k,\sigma_k^2)$ からの k 個の独立な標本である．$\bar{x}_1,\cdots,\bar{x}_k$ をこの標本の平均，$c_1,\cdots,c_k$ をすべては0でない定数とすれば，$c_1\bar{x}_1+\cdots+c_k\bar{x}_k$ は標本分布 $N\left(\sum_{i=1}^{k}c_i\mu_i,\ \sum_{i=1}^{k}\dfrac{c_i^2\sigma_i^2}{n_i}\right)$ を持つ．

これは特性関数を用いると簡単に確かめられる．読者自身で試してみよ．特に，$c_1=1$, $c_2=-1$, $c_3=\cdots=c_k=0$ とすれば，次の系が得られる．

8.3.4 a　2つの独立な正規分布 $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$ からの標本 $(x_{11},\cdots,x_{1n_1})$, $(x_{21},\cdots,x_{2n_2})$ の標本平均の差 $\bar{x}_1-\bar{x}_2$ の分布は $N(\mu_1-\mu_2,\sigma_1^2/n_1+\sigma_2^2/n_2)$ である．

8.4　正規分布からの標本の2次形式

8.3.3 d, 8.3.3 e, 8.3.4, 8.3.4 a では，正規分布からの標本和，標本平均に関する本質的な性質を述べた．この節では，1次元正規分布からの標本について，2乗和，分散，2次形式を考察しよう．k 次元の場合は，やや複雑になるが，これについては，第18章で，正規多変量解析として述べる．まず，標本2乗和に関する基本的な性質から述べよう．

8.4.1　$(x_1,\cdots,x_n)$ が正規分布 $N(\mu,\sigma^2)$ からの標本のとき，標本2乗和 $\dfrac{1}{\sigma^2}\sum_{\xi=1}^{n}(x_\xi-\mu)^2$ はカイ2乗分布 $C(n)$ に従う．

これを確かめるには，正規分布 $N(\mu,\sigma^2)$ の p.d.f. を (8.1.3) に代入したとき，標本 $(x_1,\cdots,x_n)$ の p.d.f. の指数項に $\dfrac{1}{\sigma^2}\sum_{\xi=1}^{n}(x_\xi-\mu)^2$ が現われることを示せば十分で

ある．**7.8.2** より，この2次形式はカイ2乗分布 $C(n)$ を持つからである．この結果は Helmert (1876 *a*) が最初に求めた．

(8.2.6) で定義されているように，標本分散 s^2 は $\sum_{\xi=1}^{n}(x_\xi-\bar{x})^2$ である．ただし $\bar{x}$ は標本平均．この s^2 については次が成り立つ．

8.4.2　$(x_1,\cdots,x_n)$ が正規分布 $N(\mu,\sigma^2)$ からの確率のとき，$\sqrt{n}(\bar{x}-\mu)/\sigma$ と $(n-1)s^2/\sigma^2$ は統計的に独立で，それぞれ正規分布 $N(0,1)$，カイ2乗分布 $C(n-1)$ に従う．

8.3.3 d より，$\bar{x}$ は分布 $N(\mu,\sigma^2/n)$ を持つ．これは $\sqrt{n}(\bar{x}-\mu)/\sigma$ が $N(0,1)$ に従うことと同値である．**8.4.2** を示すには $\sqrt{n}(\bar{x}-\mu)/\sigma$ と $(n-1)s^2/\sigma^2$ が独立で，$(n-1)s^2/\sigma^2$ がカイ2乗分布 $C(n-1)$ に従うことをいえばよい．まず，確率変数 $(\sqrt{n}(\bar{x}-\mu)/\sigma,\ (n-1)s^2/\sigma^2)$ の特性関数

(8.4.1)

$$\varphi(t_1,t_2)=\mathscr{E}\left[\exp\left(\mathrm{i}t_1(\bar{x}-\mu)\frac{\sqrt{n}}{\sigma}+\mathrm{i}t_2\frac{(n-1)s^2}{\sigma^2}\right)\right]$$
$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\int_{R_n}\exp\left\{-\frac{1}{2}\left[Q_0(x_1,\cdots,x_n)-2\mathrm{i}t_1(\bar{x}-\mu)\frac{\sqrt{n}}{\sigma}\right]\right\}dx_1\cdots dx_n$$

を考える．ただし

(8.4.2)　$$Q_0(x_1,\cdots,x_n)=\frac{1}{\sigma^2}\sum_{\xi=1}^{n}(x_\xi-\mu)^2-\frac{2\mathrm{i}t_2}{\sigma^2}\sum_{=1}^{n}(x_\xi-\bar{x})^2.$$

(8.4.3)　$$\tau^{\xi\xi}=\frac{1}{\sigma^2}\left[1-2\mathrm{i}t_2\left(1-\frac{1}{n}\right)\right],\qquad \xi=1,\cdots,n$$
$$\tau^{\xi\eta}=\tau^{\eta\xi}=\frac{2\mathrm{i}t_2}{n\sigma^2},\qquad \xi\neq\eta=1,\cdots,n$$

と置くと

(8.4.4)　$$Q_0(x_1,\cdots,x_n)=\sum_{\xi,\eta=1}^{n}\tau^{\xi\eta}(x_\xi-\mu)(x_\eta-\mu)$$

になるから，(7.4.15) を用いると

(8.4.5)　$$Q_0(x_1,\cdots,x_n)-\frac{2\mathrm{i}t_1}{\sigma}(\bar{x}-\mu)\sqrt{n}=Q_0'(x_1,\cdots,x_n)+\left(\sum_{\xi,\eta=1}^{n}\tau_{\xi\eta}\right)\frac{t_1^2}{n\sigma^2}.$$

ここで

(8.4.6)　$$Q_0'(x_1,\cdots,x_n)=Q_0(x_1-\mathrm{i}g_1,\cdots,x_n-\mathrm{i}g_n)$$

で，

$$\|\tau_{\xi\eta}\|=\|\tau^{\xi\eta}\|^{-1};\qquad g_\xi=\frac{t_1}{\sqrt{n}\,\sigma}\sum_{\eta=1}^{n}\tau_{\xi\eta},\qquad \xi=1,\cdots,n.$$

行列 $\|\tau^{\xi\eta}\|$ は

(8.4.7)　$$\begin{Vmatrix} a & b & \cdot & \cdot & \cdot & b \\ b & a & \cdot & \cdot & \cdot & b \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ b & b & \cdot & \cdot & \cdot & a \end{Vmatrix}$$

なる形をしていることにより，行列式は $(a-b)^{n-1}[a+(n-1)b]$ である．この逆行列も，

(8.4.8)　$$\begin{Vmatrix} A & B & \cdot & \cdot & \cdot & B \\ B & A & \cdot & \cdot & \cdot & B \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ B & B & \cdot & \cdot & \cdot & A \end{Vmatrix}$$

なる形になる．ただし

(8.4.9)　$$A=\frac{a+(n-2)b}{(a-b)[a+(n-1)b]},\qquad B=\frac{-b}{(a-b)[a+(n-1)b]}.$$

ゆえに，

(8.4.10)　$$\tau_{\xi\xi}=\frac{\sigma^2\left(1-2\mathrm{i}\frac{t_2}{n}\right)}{(1-2\mathrm{i}t_2)},\qquad \xi=1,\cdots,n$$
$$\tau_{\xi\eta}=\frac{-\sigma^2\left(2\mathrm{i}\frac{t_2}{n}\right)}{(1-2\mathrm{i}t_2)},\qquad \xi\neq\eta=1,\cdots,n$$

で，

(8.4.11)　$$\sum_{\xi,\eta=1}^{n}\tau_{\xi\eta}=n\sigma^2$$

となる．(8.4.11) を (8.4.5) に代入し，(8.4.1) の積分において，(8.4.5) を用いれば，

(8.4.12)　$$\varphi(t_1,t_2)=e^{-\frac{1}{2}t_1^2}\cdot\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n\int_{R_n}e^{-\frac{1}{2}Q_0'(x_1,\cdots,x_n)}dx_1\cdots dx_n.$$

(7.4.16) で $H=1$ を示したように，(8.4.12) の積分値は $\sqrt{(2\pi)^n}\,|\tau^{\xi\eta}|^{-\frac{1}{2}}$ である．

$$|\tau^{\xi\eta}|=\frac{1}{\sigma^{2n}}(1-2\mathrm{i}t_2)^{n-1}$$

より，結局

$$\varphi(t_1, t_2) = \varphi_1(t_1)\cdot\varphi_2(t_2) \tag{8.4.13}$$

を得る．ただし

$$\varphi_1(t_1) = e^{-\frac{1}{2}t_1^2} \tag{8.4.14}$$

$$\varphi_2(t_2) = (1 - 2it_2)^{-\frac{1}{2}(n-1)}. \tag{8.4.15}$$

これで $\varphi(t_1, t_2)$ は (8.4.13) で示されたように φ_1 と φ_2 に分解できたので，**5.3.1** から，$\sqrt{n}(\bar{x} - \mu)/\sigma$ と $(n-1)s^2/\sigma^2$ は独立で，その特性関数はそれぞれ (8.4.14), (8.4.15) を持つ．(8.4.14), (8.4.15) はそれぞれ分布 $N(0, 1)$, $C(n-1)$ の特性関数に他ならない [**7.2.1**, (7.8.2) を参照せよ]．特性関数により分布が一意に定まるから **8.4.2** が成立する．

8.4.2 と 7.8(b) 節 より，次の重要な結果が成り立つ．

8.4.3　$(x_1, \cdots, x_n)$ は $N(\mu, \sigma^2)$ からの標本とする．このとき

$$t = \sqrt{n}(\bar{x} - \mu)/s \tag{8.4.16}$$

は“スチューデント”分布 $S(n-1)$ を持つ．

なぜなら，(7.8.3) において，$u = \sqrt{n}(\bar{x} - \mu)/\sigma$, $v = (n-1)s^2/\sigma^2$ と置くと，$t = \dfrac{u}{\sqrt{v/(n-1)}} = \sqrt{n}(\bar{x} - \mu)/s$ となる．しかも，$N(\mu, \sigma^2)$ からの標本のとき，**8.4.2** より u と v は独立で，それぞれ $N(0, 1)$, $C(n-1)$ を持つ．よって，**7.8.3** から **8.4.3** が得られる．

注意　**8.4.3** は W. S. Gosset (1908) が“スチューデント”の名のもとに最初に得た．後に，この結果は **8.4.2** とともに，Fisher (1926 *a*) により確認された．フィッシャーは (8.4.16) で定義された比と p. d. f. (7.8.4) を持つ一般的な分布との両方に“スチューデント”という名前をつけた．v が自由度 $(n-1)$ のカイ2乗分布を持つという事実は Helmert (1876 *b*) により，はじめて示された．

8.4.2 の逆も，また，正しいことに注意すべきである．すなわち，$\bar{x}$, s^2 がそれぞれ p. d. f. $f(x)$ からの大きさ n の標本平均，標本分散で，互いに独立であるならば，$f(x)$ は正規分布の p. d. f. であるといえる．この結果は Geary (1936) により最初に考察されたが，後に Kawata と Sakamoto (1949), Lukacs (1942) が得た．問題 8.33 を参照せよ．

Cramér (1936) は，x_1, x_2 が独立で，和 $x_1 + x_2$ が正規分布に従う確率変数のとき，

x_1, x_2 はともに正規分布を持つことを示した．これは，帰納法により，2個以上の確率変数のときにも拡張される．

標本を分布 $N(\mu, \sigma^2)$ からとったとき，(8.4.14) で定義された t および t の分布は，ともに σ^2 に無関係であるという意味において，後の章で見るように，統計的推測論の見地から大変重要である．

再度，$N(\mu, \sigma^2)$ からの標本 $(x_1, \cdots, x_n)$ の p. d. f. は2乗和 $Q = \dfrac{1}{\sigma^2}\sum_{\xi=1}^{n}(x_\xi - \mu)^2$ を指数項に持ち，**8.4.1** より，Q はカイ2乗分布 $C(n)$ に従うことを思い出そう．Q は次のように分解される．

$$Q = Q_1 + Q_2 \tag{8.4.17}$$

$$Q_1 = \frac{1}{\sigma^2}\sum_{\xi=1}^{n}(x_\xi - \bar{x})^2, \qquad Q_2 = \frac{n}{\sigma^2}(\bar{x} - \mu)^2. \tag{8.4.18}$$

$\sqrt{n}(\bar{x} - \mu)/\sigma$ は $N(0, 1)$ に従うから，**7.8.2 a** より $[\sqrt{n}(\bar{x} - \mu)/\sigma]^2$, すなわち Q_2 はカイ2乗分布 $C(1)$ に従う．Q_1 と $\sqrt{n}(\bar{x} - \mu)/\sigma$ の独立性から Q_1 と Q_2 も独立になる．このようにカイ2乗分布 $C(n)$ に従う Q は，独立で，それぞれカイ2乗分布 $C(n-1)$, $C(1)$ に従う Q_1, Q_2 に分解できる．よって次の問題が考えられる．正規分布からの標本 $(x_1, \cdots, x_n)$ のとき，標本 p. d. f. の指数項の2乗和がいくつかの互いに独立なカイ2乗分布を持つ要素に分解可能なのはどのような条件のときか？ これに対する答は Cochran (1934) による次の定理に示されている．正規分布 $N(0, 1)$ からの標本について述べれば十分である．

8.4.4　$(x_1, \cdots, x_n)$ は正規分布 $N(0, 1)$ からの標本で，$\sum_{\xi=1}^{n} x_\xi^2 \equiv \sum_{i=1}^{k} Q_i$ とする．ただし Q_i は階数 n_i の $x_1, \cdots, x_n$ の非負の2次形式である．このとき，各 Q_i がカイ2乗分布 $C(n_i)$, $i = 1, \cdots, x$ を持ち，互いに独立であるための必要十分条件は $n = \sum_{i=1}^{k} n_i$ である．

まず，必要性を示そう．Q_i が独立でカイ2乗分布 $C(n_i)$, $i = 1, \cdots, k$ に従うならば，$\sum_{i=1}^{k} n_i = n$ が成り立つことをいう．カイ2乗分布の再生性より，$Q_1 + \cdots + Q_k$ はカイ2乗分布 $C(n_1 + \cdots + n_k)$ を持つ．しかも $Q_1 + \cdots + Q_k \equiv \sum_{\xi=1}^{n} x_\xi^2$ で，$(x_1, \cdots, x_n)$ が $N(0, 1)$ からの標本であるので，**8.4.1** を用いると，$\sum_{\xi=1}^{n} x_\xi^2$ はカイ2乗分布 $C(n)$ を持つ．したがって，これは $C(n_1 + \cdots + n_k)$ と一致する．ゆえに，$\sum_{i=1}^{k} n_i = n$ が成り立つ．

十分性を示すために $n_1+\cdots+n_k=n$ を仮定しよう．次を満たす正則線形変換の存在を示さねばならない．

(8.4.19)
$$x_\xi=\sum_{\eta=1}^n b_{\xi\eta}y_\eta,\qquad \xi=1,\cdots,n$$

が $Q_1,\cdots,Q_k$ を次のように変換する．

(8.4.20)
$$\begin{aligned}Q_1&=y_1^2+\cdots+y_{n_1}^2\\ Q_2&=y_{n_1+1}^2+\cdots+y_{n_1+n_2}^2\\ &\cdots\cdots\cdots\cdots\cdots\\ Q_k&=y_{n_1+\cdots+n_{k-1}+1}^2+\cdots+y_{n_1+\cdots+n_k}^2.\end{aligned}$$

ただし $n_1+\cdots+n_k=n$ で，$y_1,\cdots,y_n$ はすべて $N(0,1)$ に従う独立な確率変数である．これが成り立てば，$(x_1,\cdots,x_n)$ の p.e. を $N(0,1)$ からの標本 $(y_1,\cdots,y_n)$ の p.e. に変換できる．よって $Q_1,\cdots,Q_k$ はそれぞれ $n_1,\cdots,n_k$ 個の互いに素で，すべてが $N(0,1)$ を持つ独立な確率変数の2乗和である．ゆえに，$Q_1,\cdots,Q_k$ は互いに独立で，**8.4.1** より，カイ2乗分布 $C(n_1),\cdots,C(n_k)$ に従うことになる．

いま，Q_1 は階数 n_1 の行列の非負の2次形式だから，n_1 個の1次独立な $x_1,\cdots,x_n$ の線形結合 $y_1,\cdots,y_{n_1}$ が存在する［たとえば，Bôcher (1907) あるいは Birkhoff と MacLane (1953) を参照せよ］．すなわち

(8.4.21)
$$y_{\xi_1}=\sum_{\eta=1}^n b^{\xi_1\eta}x_\eta,\qquad \xi_1=1,\cdots,n_1$$

で，

(8.4.22)
$$Q_1=y_1^2+\cdots+y_{n_1}^2.$$

同様に $Q_2,\cdots,Q_k$ に対しても，それぞれ $n_2,\cdots,n_k$ 個の1次独立な $x_1,\cdots,x_n$ の線形結合が存在する．

(8.4.23)
$$\begin{aligned}y_{\xi_2}&=\sum_{\eta=1}^n b^{\xi_2\eta}x_\eta,\qquad \xi_2=n_1+1,\cdots,n_1+n_2\\ &\ \ \vdots\\ y_{\xi_k}&=\sum_{\eta=1}^n b^{\xi_k\eta}x_\eta,\qquad \xi_k=n_1+\cdots+n_{k-1}+1,\cdots,n_1+\cdots+n_{k-1}+n_k.\end{aligned}$$

1つの Q に関する y については，1次従属性は成り立っていない．任意の Q を固定したとき，Q と異なる他の2つないしはそれ以上の Q に関する y についても1次従属性が成り立たないことを示せばよい．いま，1次従属性が成り立つと仮定すれば，ある θ を Q^* としたとき，Q^* に関する任意の y は，Q^* 以外の Q の1つあるいは複数個の Q からの y の線形結合として表わせる．すなわち，$Q_1+\cdots+Q_k$ は y の高々 $n-1$ 個の非負の2次形式として表わせることである．各 y は $x_1,\cdots,x_n$ の線形結合だから，$Q_1+\cdots+Q_k$ は階数が高々 $(n-1)$ の行列に関する x の非負の2次形式になる．しかし $Q_1+\cdots+Q_k$ は，$\sum_{\xi=1}^n x_\xi^2$ に等しくなっており，この階数は n である．ゆえに，任意の Q，たとえば Q^*，に関する y は，Q^* 以外の1つあるいはそれ以上の Q に関する y の線形結合として表わすことはできない．よって，行列

$$\|b^{\xi\eta}\|,\qquad \xi,\eta=1,\cdots,n$$

は正則であることが示された．変換

(8.4.24)
$$y_\xi=\sum_{\eta=1}^n b^{\xi\eta}x_\eta,\qquad \xi=1,\cdots,n$$

は正則で，唯一の逆変換

(8.4.25)
$$x_\eta=\sum_{\xi=1}^n b_{\xi\eta}y_\xi,\qquad \eta=1,\cdots,n$$

を持つ．よって，$\sum_{\xi=1}^n x_\xi^2$ すなわち $Q_1+\cdots+Q_k$ は

$$(y_1^2+\cdots+y_{n_1}^2)+(y_{n_1+1}^2+\cdots+y_{n_1+n_2}^2)+\cdots+(y_{n_1+\cdots+n_{k-1}+1}^2+\cdots+y_n^2)$$

に変換される．よって，この変換は直交であることから，y はすべて $N(0,1)$ を持つ独立な確率変数になる．ゆえに $Q_1,\cdots,Q_k$ はそれぞれ $(y_1^2+\cdots+y_{n_1}^2),\cdots,(y_{n_1+\cdots+n_{k-1}+1}^2+y_n^2)$ と同じ分布を持つ．したがって，$(y_1^2+\cdots+y_{n_1}^2),\cdots,(y_{n_1+\cdots+n_{k-1}+1}^2+\cdots+y_n^2)$ はそれぞれカイ2乗分布 $C(n_1),\cdots,C(n_k)$ に従う確率変数である．これで十分性が示され，**8.4.4** の証明が終わる．

コクランの定理は，分散分析や，第10章で見るように，回帰理論において適用されている．

8.5　有限母集団からの標本

(a)　1次元の場合

8.1 節で与えた標本$(x_1,\cdots,x_n)$において，各要素は同じ c.d.f. $F(x)$ を持つ独立な確率変数として定義された．この定義に基づく標本論は c.d.f. $F(x)$ を持つ無限母集団か

らの単純確率標本論と呼ばれている．これは次のように解釈される．標本要素 $x_1, \cdots, x_n$ を母集団と同じ分布を持つ独立な確率変数列とみなせば，$x_1, \cdots, x_n$ は母集団から逐次ぬき出した結果を表わしていて，ぬき取りを行なっても母集団の分布が変わらないものと考えられる．特に断わらない限り，"大きさ n の標本" といえば，常に無限母集団からの標本を意味するものとする．

有限母集団 π_N からの単純確率標本に関する理論も展開されている．ここでは π_N は N 個の要素 $o_1, \cdots, o_N$ からなる母集団とする．π_N を標本点 $o_1, \cdots, o_N$ を持つ標本空間と考え，$x(o)$ を各標本点で定義された確率変数，すなわち $x(o_t) = x_{ot}$，$t = 1, \cdots, N$ とすれば，$x(o)$ は $o_1, \cdots, o_N$ を R_1 の点 $x_{o1}, \cdots, x_{oN}$ に写す写像となる．$x_{o1}, \cdots, x_{oN}$ は π_N の N 個の要素の x による値*) とみなせる．一般性を失うことなく，便宜上，$x_{o1} \leqq \cdots \leqq x_{oN}$ となるように π_N の要素にラベルをつけておこう．$o_1, \cdots, o_N$ に等確率 $\dfrac{1}{N}$ を与え，確率変数 $x(o)$ を x，p.f. を $p(x)$ とすると，x の質点は $x_{o1}, \cdots, x_{oN}$ となり，各点で

(8.5.1)
$$p(x) = \frac{1}{N}$$

となる．$x(o)$ が π_N の複数個の点で同じ値をとるとき，$p(x)$ は $\dfrac{1}{N}\times$(点の個数) になる．すなわち，$x(o) = x'$ が π_N の r 個の点で成り立てば，$p(x') = r/N$ である．以後，x を π_N から "無作為にぬき取られた" 要素の x 値を示す確率変数と考える．このように定義された $p(x)$ を有限母集団の p.f. と呼ぶ．

いま，$(o_{\tau_1}, \cdots, o_{\tau_n})$ を π_N の N 個の要素のうち，n 個を取り出したときのある置換とする．$N!/(N-n)!$ 通りの n 置換がある．この n 置換を標本空間 R の標本点と考える．$(o_{\tau_1}, \cdots, o_{\tau_n})$ を R の標本点 e とするとき，$x_1(e), \cdots, x_n(e)$ を $x_1(e) = x_{o_{\tau_1}}, \cdots, x_n(e) = x_{o_{\tau_n}}$ なる確率変数とする．$(x_1(e), \cdots, x_n(e))$ を $(x_1, \cdots, x_n)$ と書くと，この n 次元確率変数は有限母集団 π_N からの大きさ n の標本と考えられる．x_1 は π_N から無作為にぬき取られた最初の要素の x 値を示す確率変数で，x_2 は重複を許さずに2番目にぬき取られたときの x 値を示す．以下 $x_3, \cdots, x_{n-1}$ も同じく，x_n は π_N から重複を許さずに取り出された n 番目の要素の x 値を示している．

明らかに，$(x_1, \cdots, x_n)$ は離散型 n 次元確率変数である．R のすべての標本点 e に等確率，すなわち

*) 以下，簡単に "x 値" という．(訳注)

(8.5.2)
$$\frac{(N-n)!}{N!}$$

を与えれば，$(x_1, \cdots, x_n)$ の p.f. $p(x_1, \cdots, x_n)$ は $(N-n)!/N!$ の "点の個数倍" である．

実際は，以下の議論において見るように，π_N の N 個の要素の x 値はすべて異なっていて，$x_{o1} < \cdots < x_{oN}$ となるような場合を考えると便利である．このとき確率変数 x_ξ $(\xi = 1, \cdots, n)$ の質点は $x_{o1}, \cdots, x_{oN}$ になる．$R_1^{(\xi)}$ を確率変数 x_ξ に対応する実数軸，$E_1^{(\xi)}$ を $R_1^{(\xi)}$ の質点 $x_{o1}, \cdots, x_{oN}$ の集合とする．E_n を直積 $E_1^{(1)} \times \cdots \times E_1^{(n)}$，$R_n$ を直積 $R_1^{(1)} \times \cdots \times R_1^{(n)}$ とすれば，$(x_1, \cdots, x_n)$ の質点は E_n の部分集合 E'_n を構成する．ただし E'_n はどの2つの座標も等しくない E_n の点全体である．E'_n の点はちょうど $N!/(N-n)!$ 個からなり，E'_n の点と R の e 点と1対1の対応がつく．すなわち，n 次元確率変数 $(x_1(e), \cdots, x_n(e))$ は R の e 点を R_n における E'_n の点へ写像する．よって，標本空間 R_n の各点で定義される $(x_1, \cdots, x_n)$ の p.f. は次のようになる．

(8.5.3)
$$p(x_1, \cdots, x_n) = \begin{cases} \dfrac{(N-n)!}{N!}, & x_1 \neq \cdots \neq x_n \\ 0, & \text{その他.} \end{cases}$$

要素 $o_{\tau_1}, \cdots, o_{\tau_n}$ に対応する $(x_1, \cdots, x_n)$ の質点は $(x_{o_{\tau_1}}, \cdots, x_{o_{\tau_n}})$ である．(8.5.3) における $(x_1, \ldots, x_{n-1})$ の周辺 p.f.，すなわち，$p(x_1, \cdots, x_n)$ の x_n に関する和は

$$p_{1\cdots(n-1)}(x_1, \cdots, x_{n-1}) = (N-n+1)\frac{(N-n)!}{N!} = \frac{(N-n+1)!}{N!}$$

で与えられる．ただし $(x_1, \cdots, x_{n-1})$ は $E_1^{(1)} \times \cdots \times E_1^{(n-1)}$ の $(n-1)$ のすべての座標が異なっている点である．その他の $E_1^{(1)} \times \cdots \times E_1^{(n-1)}$ の点に対しては $p_{1\cdots(n-1)}(x_1, \cdots, x_{n-1})$ はすべて0である．一般に，$(x_{\xi_1}, \cdots, x_{\xi_r})$ の周辺 p.f. は

(8.5.4)
$$p_{\xi_1\cdots\xi_r}(x_{\xi_1}, \cdots, x_{\xi_r}) = \begin{cases} \dfrac{(N-r)!}{N!}, & x_{\xi_1} \neq \cdots \neq x_{\xi_r} \\ 0, & \text{その他} \end{cases}$$

になる．$(x_1, \cdots, x_n)$ の唯一の成分，たとえば，x_ξ の p.f. は (8.5.1) で表わされる．すなわち

(8.5.5)
$$p_\xi(x_\xi) = \begin{cases} \dfrac{1}{N}, & x_\xi = x_{o1}, \cdots, x_{oN} \\ 0, & \text{その他.} \end{cases}$$

このように $(x_1, \cdots, x_n)$ のどの成分も母集団と同じ分布を持つ．

有限母集団 π_N の平均 μ_N，分散 σ_N^2 は

(8.5.6)
$$\mu_N = \frac{1}{N}\sum_{t=1}^{N} x_{ot}, \qquad \sigma_N^2 = \frac{1}{N-1}\sum_{t=1}^{N}(x_{ot}-\mu_N)^2$$

で定義される．μ_N，σ_N の N を落として，簡単に μ，σ^2 と書くことにしよう．$\xi=1, \cdots, n$ に対して，

(8.5.7)
$$\mathscr{E}(x_\xi)=\mu, \qquad \sigma^2(x_\xi)=\frac{N-1}{N}\sigma^2$$

が成り立つことは明らかである．$r=2$ の場合，(8.5.4) の周辺 p.f. を求めると便利である．$(x_1, \cdots, x_n)$ の任意の2つの成分 (x_ξ, x_η) の p.f. は

(8.5.8)
$$p_{\xi\eta}(x_\xi, x_\eta)=\frac{1}{N(N-1)}, \qquad (x_\xi, x_\eta)=(x_{ot}, x_{ot'}).$$
$$t \neq t'=1, \cdots, N.$$

x_ξ，x_η の共分散は

(8.5.9)
$$\begin{aligned}\mathrm{cov}(x_\xi, x_\eta) &= \sum_{x_\xi, x_\eta}(x_\xi-\mu)(x_\eta-\mu)p_{\xi\eta}(x_\xi, x_\eta)\\ &= \sum_{t\neq t'=1}^{N}\frac{(x_{ot}-\mu)(x_{ot'}-\mu)}{N(N-1)}\\ &= -\sum_{t=1}^{N}\frac{(x_{ot}-\mu)^2}{N(N-1)}+\frac{1}{N(N-1)}\left[\sum_{t=1}^{N}(x_{ot}-\mu)\right]^2.\end{aligned}$$

(8.5.6) より

(8.5.10)
$$\mathrm{cov}(x_\xi, x_\eta)=-\frac{\sigma^2}{N}.$$

相関係数は

(8.5.11)
$$\rho(x_\xi, x_\eta)=-\frac{1}{N-1}.$$

無限母集団からの標本の場合には，(8.2.4) から $\mathrm{cov}(x_\xi, x_\eta)=0$ となることに注意せよ．まとめると

8.5.1　$(x_1, \cdots, x_n)$ が有限母集団 π_N からの大きさ n の標本のとき

$$\mathscr{E}(x_\xi)=\mu, \qquad \sigma^2(x_\xi)=\frac{N-1}{N}\sigma^2, \qquad \xi=1, \cdots, n$$
$$\mathrm{cov}(x_\xi, x_\eta)=-\frac{\sigma^2}{N}, \qquad \rho(x_\xi, x_\eta)=-\frac{1}{N-1}, \qquad \xi \neq \eta=1, \cdots, n.$$

ただし，μ，σ^2 は (8.5.6) で定義されている．

有限母集団 π_N からの標本 $(x_1, \cdots, x_n)$ の任意の関数 $g(x_1, \cdots, x_n)$ の標本論に関する問題は，主に $g(x_1, \cdots, x_n)$ の分布を求めることである．ただし $(x_1, \cdots, x_n)$ は (8.5.3) を p.f. に持つ確率変数である．次の結果は $x_{o1}<\cdots<x_{oN}$ のいくつかに等号がはいっても，少し変更すれば成り立つことに注意せよ．

(b)　標本平均と標本分散の期待値

まず，$(x_1, \cdots, x_n)$ の特殊な関数として $\bar{x}$ と s^2 を考えよう．$\bar{x}=\frac{1}{n}\sum_{\xi=1}^{n}x_\xi$ より

(8.5.12)
$$\mathscr{E}(\bar{x})=\mathscr{E}\left(\frac{1}{n}\sum_\xi x_\xi\right)=\frac{1}{n}\sum_\xi \mathscr{E}(x_\xi)$$
$$\mathscr{E}(x_\xi)=\mu, \qquad \xi=1, \cdots, n.$$

よって

(8.5.13)
$$\mathscr{E}(\bar{x})=\mu.$$

ゆえに，$\bar{x}$ は μ の不偏推定量である．

$\sigma^2(\bar{x})$ に対しては，

$$\begin{aligned}\sigma^2(\bar{x}) &= \mathscr{E}(\bar{x}-\mu)^2\\ &= \mathscr{E}\left[\frac{1}{n}\sum_\xi(x_\xi-\mu)\right]^2=\frac{1}{n^2}\sum_\xi \mathscr{E}(x_\xi-\mu)^2+\frac{1}{n^2}\sum_{\xi\neq\eta}\mathscr{E}(x_\xi-\mu)(x_\eta-\mu).\end{aligned}$$

$\mathscr{E}(x_\xi-\mu)^2=\frac{N-1}{N}\sigma^2$，$\xi=1, \cdots, n$ で，$\xi \neq \eta$ のとき

$$\mathscr{E}(x_\xi-\mu)(x_\eta-\mu)=\mathrm{cov}(x_\xi, x_\eta)=-\frac{\sigma^2}{N}$$

より

(8.5.14)
$$\sigma^2(\bar{x})=\left(\frac{1}{n}-\frac{1}{N}\right)\sigma^2.$$

s^2 の期待値を考える．

$$\begin{aligned}s^2 &= \frac{1}{n-1}\sum_\xi(x_\xi-\bar{x})^2\\ &= \frac{1}{n}\sum_\xi(x_\xi-\mu)^2-\frac{1}{n(n-1)}\sum_{\xi\neq\eta}(x_\xi-\mu)(x_\eta-\mu)\end{aligned}$$

の期待値をとると，**8.5.1** が使えて

(8.5.15)
$$\mathscr{E}(s^2)=\sigma^2$$

となり，s^2 は σ^2 に対する不偏推定量である．これらをまとめると

8.5.2 $(x_1, \cdots, x_n)$ が平均 μ, 分散 σ^2 を持つ有限母集団 π_N からの標本のとき

$$\mathscr{E}(\bar{x}) = \mu, \qquad \sigma^2(\bar{x}) = \left(\frac{1}{n} - \frac{1}{N}\right)\sigma^2, \qquad \mathscr{E}(s^2) = \sigma^2$$

が成り立つ.

標本和 z の平均と分散は

$$\mathscr{E}(z) = n\mu, \qquad \sigma^2(z) = n^2\left(\frac{1}{n} - \frac{1}{N}\right)\sigma^2 \tag{8.5.16}$$

で与えられる.

最後に, (8.5.6) で定義された $x_{o1}, \cdots, x_{oN}$ の関数 μ_N と σ_N^2 が, $N \longrightarrow \infty$ のとき, ともに有限な極限値 μ, σ^2 に近づくならば, **8.5.2** の $\mathscr{E}(\bar{x})$, $\sigma^2(\bar{x})$ は, $N \longrightarrow \infty$ のとき, **8.2.1** で与えられた無限母集団からの標本に対する $\mathscr{E}(\bar{x})$ と $\sigma^2(\bar{x})$ に近づくことに注意せよ.

(c) 有限母集団 π_N からの標本対称関数の期待値

8.5 (b) 節における基本的な考え方を拡張すると, 標本のある種の対称関数の期待値が得られる. $(x_1, \cdots, x_n)$ を有限母集団 π_N からの標本とし, $r < n$ なるすべての $\xi_1, \cdots, \xi_r$ に対して, 関数 $g(x_{\xi_1}, \cdots, x_{\xi_r})$ を考えよう. (8.5.4) で与えられた $(x_{\xi_1}, \cdots, x_{\xi_r})$ の p.f. の型より

$$\mathscr{E}[g(x_{\xi_1}, \cdots, x_{\xi_r})] = \frac{(N-r)!}{N!} \sum\nolimits_t' g(x_{ot_1}, \cdots, x_{ot_r}) \tag{8.5.17}$$

になる. ただし $\sum_t'$ は, $t_1, \cdots, t_r$ が 1 から N までのすべての異なる値をとる和を表わす. さて, $(x_1, \cdots, x_n)$ の対称関数 Q を

$$Q(x_1, \cdots, x_n) = \frac{(n-r)!}{n!} \sum\nolimits_\xi' g(x_{\xi_1}, \cdots, x_{\xi_r}) \tag{8.5.18}$$

としよう. $\sum_\xi'$ は $t_1, \cdots, t_r$ に対する $\sum_t'$ と同様に定義される和を表わす. この $Q(x_1, \cdots, x_n)$ の期待値は

$$\mathscr{E}(Q(x_1, \cdots, x_n)) = \frac{(n-r)!}{n!} \sum\nolimits_\xi' \mathscr{E}[g(x_{\xi_1}, \cdots, x_{\xi_r})]. \tag{8.5.19}$$

ここに, $\mathscr{E}[g(x_{\xi_1}, \cdots, x_{\xi_r})]$ は $\xi_1 \neq \xi_2 \neq \cdots \neq \xi_r$ なるすべての $(\xi_1, \cdots, \xi_r)$ について, (8.5.17) の右辺で与えられる同じ値を持つ. このような $(\xi_1, \cdots, \xi_r)$ は $n!/(n-r)!$ 個ある. よって

$$\mathscr{E}(Q(x_1, \cdots, x_n)) = \frac{(N-r)!}{N!} \sum\nolimits_t' g(x_{ot_1}, \cdots, x_{ot_r}) = Q(x_{o1}, \cdots, x_{oN}). \tag{8.5.20}$$

すなわち, $Q(x_1, \cdots, x_n)$ は $Q(x_{o1}, \cdots, x_{oN})$ に対する不偏推定量である. よって, 次の結果が得られた.

8.5.3 $(x_1, \cdots, x_n)$ が有限母集団 π_N からの標本で, $Q(x_1, \cdots, x_n)$ は (8.5.18) で定義された $(x_1, \cdots, x_n)$ の対称関数のとき, $Q(x_1, \cdots, x_n)$ は $Q(x_{o1}, \cdots, x_{oN})$ の不偏推定量である.

さらに, **3.2.1**, **3.2.2** により, 次が成立つ.

8.5.4 $Q_1(x_1, \cdots, x_n), \cdots, Q_s(x_1, \cdots, x_n)$ が有限母集団 π_N からの (8.5.18) なる型の任意の対称関数であるならば, 任意定数 $d_1, \cdots, d_s$ に対して, $d_1Q_1(x_1, \cdots, x_n) + \cdots + d_sQ_s(x_1, \cdots, x_n)$ は $d_1Q_1(x_{o1}, \cdots, x_{oN}) + \cdots + d_sQ_s(x_{o1}, \cdots, x_{oN})$ の不偏推定量である.

(8.5.18) において, $c_1, \cdots, c_r$ を 0 または正の整数として, $g(x_{\xi_1}, \cdots, x_{\xi_r}) \equiv x_{\xi_1}^{c_1} \cdots x_{\xi_r}^{c_r}$ とすれば, 対称関数の族 $\{Q(x_1, \cdots, x_n)\}$ が得られ, 標本平均, 標本分散, k 統計量, 標本モーメントの多項式などが $\{Q(x_1, \cdots, x_n)\}$ の線形結合として表わされる. この事柄と定理 **8.5.3**, **8.5.4** により, 平均, 分散, 高次のモーメント, k 統計量に関する議論が簡単になる. これは Tukey (1950, 1956 *a*) によって指摘された.

例題 $\mathscr{E}(\bar{x}^3)$ を求めよう.

いま

$$\begin{aligned}\bar{x}^3 &= \frac{1}{n^3} \sum_{\xi_1, \xi_2, \xi_3 = 1}^{n} x_{\xi_1}x_{\xi_2}x_{\xi_3} = \frac{1}{n^3}\left[\sum_{\xi_1=1}^{n} x_{\xi_1}^3 + 3\sum_{\xi_1 \neq \xi_2} x_{\xi_1}^2 x_{\xi_2} + \sum_{\xi_1 \neq \xi_2 \neq \xi_3} x_{\xi_1}x_{\xi_2}x_{\xi_3}\right] \\ &= \frac{1}{n^2} Q_1(x_1, \cdots, x_n) + \frac{3(n-1)}{n^2} Q_2(x_1, \cdots, x_n) \\ &\qquad + \frac{(n-1)(n-2)}{n^2} Q_3(x_1, \cdots, x_n).\end{aligned}$$

ただし

$$Q_1 = \frac{1}{n}\sum_{\xi_1=1}^{n} x_{\xi_1}^3, \qquad Q_2 = \frac{1}{n(n-1)} \sum_{\xi_1 \neq \xi_2} x_{\xi_1}^2 x_{\xi_2},$$

$$Q_3 = \frac{1}{n(n-1)(n-2)} \sum_{\xi_1 \neq \xi_2 \neq \xi_3} x_{\xi_1}x_{\xi_2}x_{\xi_3}$$

で, いずれも (8.5.18) の型をしている. **8.5.4** を用いると

$$\mathscr{E}(\bar{x}^3) = \frac{1}{n^2} Q_1(x_{o1}, \cdots, x_{oN}) + \frac{3(n-1)}{n^2} Q_2(x_{o1}, \cdots, x_{oN}) + \frac{(n-1)(n-2)}{n^2} Q_3(x_{o1}, \cdots, x_{oN}).$$

$\mu'_r = \dfrac{1}{N} \sum_{t=1}^{N} x_{ot}^r$ とすれば

$$Q_1(x_{o1}, \cdots, x_{oN}) = \mu'_3$$

$$Q_2(x_{o1}, \cdots, x_{oN}) = \frac{1}{N-1} (N\mu'_2\mu'_1 - \mu'_3)$$

$$Q_3(x_{o1}, \cdots, x_{oN}) = \frac{1}{(N-1)(N-2)} [N^2\mu_1'^3 + 2\mu'_3 - 3N\mu'_2\mu'_1]$$

となり，$\mathscr{E}(\bar{x}^3)$ は母集団の最初の3つのモーメントの多項式で表わされる．

(d) k 次元の場合

8.5 (a) 節で用いられた $x(o)$ が k 個の成分，$x_1(o), \cdots, x_k(o)$ からなるベクトルである場合を考えよう．すなわち，π_N の各点は k 個の数値によって定まる．$x_i(o_t) = x_{oi}{}^t$ のとき，x_{oit} を o_t の x_i 値という．π_N の N 個の点に対するベクトルがすべて異なっていれば，ベクトル確率変数 $(x_1(o), \cdots, x_k(o))$ は π_N の点を R_k の N 個の点に写し，o_t のこの写像による点の座標は $(x_{o1t}, \cdots, x_{okt})$，$t = 1, \cdots, N$ になる．π_N の N 個のすべての点に等確率，すなわち，$\dfrac{1}{N}$ を与えれば，k 次元確率変数 $(x_1, \cdots, x_k)$ は p.f.

(8.5.21)
$$p(x_1, \cdots, x_k) = \frac{1}{N}$$

を持ち，質点は $(x_{o1t}, \cdots, x_{okt})$，$t = 1, \cdots, N$．この $p(x_1, \cdots, x_k)$ を有限母集団 π_N の p.f. ということにする．

さて，8.5 (a) 節と同様に，$N!/(N-n)!$ 個の n 置換からなる基本標本空間 R を考え，$(o_{\gamma_1}, \cdots, o_{\gamma_n})$ を簡単に標本点 e で表わそう．$(x_{i1}(e), \cdots, x_{in}(e)\,;\, i = 1, \cdots, k)$ を $x_{i1}(e) = x_{oi\gamma_1}, \cdots, x_{in}(e) = x_{oi\gamma_n}$ なる nk 個の確率変数とする．$(x_{i1}(e), \cdots, x_{in}(e)\,;\, i = 1, \cdots, k)$ を $(x_{i\xi}\,;\, i = 1, \cdots, k,\ \xi = 1, \cdots, n)$ で表わせば，この nk 次元確率変数は p.f. (8.5.3) を持つ k 変量有限母集団 π_N からの大きさ n の標本となる．nk 次元確率変数の標本空間 R_{nk} には $N!/(N-n)!$ 個の質点があり，各点で確率 $(N-n)!/N!$ が与えられる．$(o_{\gamma_1}, \cdots, o_{\gamma_n})$ に対応する質点は $(x_{oi\gamma_1}, \cdots, x_{oi\gamma_n},\ i = 1, \cdots, k)$ である．

このようにすると，標本つまり nk 次元確率変数 $(x_{i\xi}\,;\, i = 1, \cdots, k\,;\, \xi = 1, \cdots, n)$ の関数に関する議論を進め易くなるだろう．標本和および標本平均は次の要素を持つベクトルである．

(8.5.22)
$$z_i = \sum_{\xi=1}^{n} x_{i\xi}, \qquad \bar{x}_i = \frac{1}{n} \sum_{\xi=1}^{n} x_{i\xi}.$$

標本共分散行列 $\|s_{ij}\|$ は

(8.5.23)
$$s_{ij} = \frac{1}{n-1} \sum_{\xi=1}^{n} (x_{i\xi} - \bar{x}_i)(x_{j\xi} - \bar{x}_j)$$
$$i, j = 1, \cdots, k.$$

π_N の平均値ベクトル $(\mu_1, \cdots, \mu_k)$ を

(8.5.24)
$$\mu_i = \frac{1}{N} \sum_{t=1}^{N} x_{oit}, \qquad i = 1, \cdots, k,$$

共分散行列 $\|\sigma_{ij}\|$ を

(8.5.25)
$$\sigma_{ij} = \frac{1}{N-1} \sum_{t=1}^{N} (x_{oit} - \mu_i)(x_{ojt} - \mu_j)$$
$$i, j = 1, \cdots, k$$

で定義すると，(8.5.12) から (8.5.15) までと同様にして

(8.5.26)
$$\mathscr{E}(\bar{x}_i) = \mu_i, \qquad \sigma^2(\bar{x}_i) = \left(\frac{1}{n} - \frac{1}{N}\right)\sigma_{ii}, \qquad i = 1, \cdots, k$$
$$\mathrm{cov}(\bar{x}_i, \bar{x}_j) = \left(\frac{1}{n} - \frac{1}{N}\right)\sigma_{ij}, \qquad i \neq j = 1, \cdots, k$$
$$\mathscr{E}(s_{ij}) = \sigma_{ij}, \qquad i, j = 1, \cdots, k$$

が示される．

8.6 行列標本

(a) 2次の行列標本

u，v は，c.d.f. $F_1(u)$，$F_2(v)$ を持つ独立な確率変数であるとする．$x(u, v)$ を確率変数とし

$$\mu = \mathscr{E}(x(u, v))$$

(8.6.1)
$$\mu_{u\cdot} = \int_{-\infty}^{\infty} x(u, v)\, dF_2(v)$$

$$\mu_{\cdot v} = \int_{-\infty}^{\infty} x(u, v)\, dF_1(u)$$

とする．さらに

(8.6.2)
$$\begin{aligned} \varepsilon_{u\cdot} &= \mu_{u\cdot} - \mu \\ \varepsilon_{\cdot v} &= \mu_{\cdot v} - \mu \\ \varepsilon_{uv} &= x(u, v) - \mu_{u\cdot} - \mu_{\cdot v} + \mu \end{aligned}$$

によって，確率変数 $\varepsilon_{u\cdot}$，$\varepsilon_{\cdot v}$，ε_{uv} を定義する．明らかに平均が0で，共分散もすべて0になる．ゆえに，次の分解定理を得る．

8.6.1　u, v が独立な確率変数のとき，確率変数 $x(u, v)$ は次のように分解される．

(8.6.3)
$$x(u, v) = \mu + \varepsilon_{u\cdot} + \varepsilon_{\cdot v} + \varepsilon_{uv}$$

ただし $\varepsilon_{u\cdot}$，$\varepsilon_{\cdot v}$，ε_{uv} は平均が0，共分散も0であり，(8.6.2) で定義される．さらに

(8.6.4)
$$\sigma^2(x(u, v)) = \sigma^2(\varepsilon_{u\cdot}) + \sigma^2(\varepsilon_{\cdot v}) + \sigma^2(\varepsilon_{uv})$$

が成立する．

注意　u, v が確率変数ではなく，任意の標本空間 $R^{(1)}$, $R^{(2)}$ からの独立な標本点の場合でも，この定理の一般型（次で示す）が成立することに注意せよ．もちろん，このような空間では確率測度が存在して，$x(u, v)$ はこの直積確率測度に関して可測な確率変数である．

分散に対して，次のような簡略化された記号を用いよう．

(8.6.5)
$$\begin{aligned} \sigma^2(x(u, v)) &= \sigma^2, & \sigma^2(\varepsilon_{u\cdot}) &= \sigma^2_{\cdot 0} \\ \sigma^2(\varepsilon_{\cdot v}) &= \sigma^2_{0\cdot}, & \sigma^2(\varepsilon_{uv}) &= \sigma^2_{\cdot\cdot} \end{aligned}$$

(8.6.4) は

(8.6.4 a)
$$\sigma^2 = \sigma^2_{\cdot 0} + \sigma^2_{0\cdot} + \sigma^2_{\cdot\cdot}$$

になる．

さて，$(u_1, \cdots, u_r)$，$(v_1, \cdots, v_s)$ をそれぞれ c.d.f. $F_1(u)$，$F_2(v)$ からの独立な標本とする．したがって，2つを合わせた標本の c.d.f. は

(8.6.6)
$$\prod_{\xi=1}^{r} F_1(u_\xi) \prod_{\eta=1}^{s} F_2(v_\eta).$$

$(x(u_\xi, v_\eta)\,;\, \xi = 1, \cdots, r,\ \eta = 1, \cdots, s)$ は新しい確率変数の組で，r 行，s 列からなる

$r \times s$ の矩形配列とみなせる．これを**2次の行列標本**という．この標本は $(r+s)$ 個の確率変数，つまり，$u_1, \cdots, u_r$，$v_1, \cdots, v_s$ だけにしか依存しない．

注意　行列標本は計量心理学や実験計画法などで基本的な役割を果たす．たとえば，計量心理学においては，行はある検査を受ける被験者の母集団，列は質問の母集団を表わしている．すると $x(u_\xi, v_\eta)$ は，$u = u_\xi$ となる被験者が $v = v_\eta$ なる質問に答えたときの“得点”を示す．したがって，r 人の被験者のグループに対して，s 個の質問からなる検査を行なえば，s 個の質問に対する r 人の得点は2次の行列標本 $x(u_\xi, v_\eta)$，$\xi = 1, \cdots, r$，$\eta = 1, \cdots, s$ を構成する．Lord (1955) は計量心理学に関して，$x(u_\xi, v_\eta)$ の種々の線形関数および2次関数について標本論を展開している．

特に，実験計画法においては，列はある種の機械からなる母集団を，行はこれらの機械の操作からなる母集団を表わしている．いま，ある実験で，無作為に η 番目の機械で，ξ 番目の操作を行なったとすれば，$x(u_\xi, v_\eta)$ は η 番目の機械で，ξ 番目の操作を行ない，一定時間経過したときの結果を示していると解釈できる．実験計画への応用は 10.6, 10.7 節で詳しく議論されるであろう．

簡潔にするために，次の記号を用いる．

(8.6.7)
$$\begin{aligned} & x_{\xi\eta} = x(u_\xi, v_\eta), \qquad x'_{\xi\eta} = x_{\xi\eta} - \mu \\ & \bar{x}_{\cdot\cdot} = \frac{1}{rs}\sum_{\xi,\eta} x_{\xi\eta}, \qquad \bar{x}_{\xi\cdot} = \frac{1}{s}\sum_{\eta} x_{\xi\eta}, \qquad \bar{x}_{\cdot\eta} = \frac{1}{r}\sum_{\xi} x_{\xi\eta} \end{aligned}$$

ただし，$\xi = 1, \cdots, r$，$\eta = 1, \cdots, s$.

(8.6.3) より，各 ξ, η に対して

(8.6.8)
$$x_{\xi\eta} = \mu + \varepsilon_{\xi\cdot} + \varepsilon_{\cdot\eta} + \varepsilon_{\xi\eta}$$

と書ける．

さて，

(8.6.9)
$$\begin{aligned} S_T &= \sum_{\xi,\eta} (x_{\xi\eta} - \bar{x}_{\cdot\cdot})^2, & S_{\cdot 0} &= \sum_{\xi,\eta} (\bar{x}_{\xi\cdot} - \bar{x}_{\cdot\cdot})^2 \\ S_{0\cdot} &= \sum_{\xi,\eta} (\bar{x}_{\cdot\eta} - \bar{x}_{\cdot\cdot})^2, & S_{\cdot\cdot} &= \sum_{\xi,\eta} (x_{\xi\eta} - \bar{x}_{\xi\cdot} - \bar{x}_{\cdot\eta} + \bar{x}_{\cdot\cdot})^2 \end{aligned}$$

とすると

$$S_T = S_{\cdot 0} + S_{0\cdot} + S_{\cdot\cdot}$$

が確かめられる．$S_{\cdot 0}$，$S_{0\cdot}$，$S_{\cdot\cdot}$ をそれぞれ**全平方和** S_T の**行要素**，**列要素**，**残差要素**と呼んでいる．$\bar{x}_{\cdot\cdot}$ の期待値と分散，S_T，$S_{\cdot 0}$，$S_{0\cdot}$，$S_{\cdot\cdot}$ の期待値を求めよう．ただし (8.6.6) の c.d.f. に関する期待値である．

$$\mathscr{E}(x_{\xi\eta}) = \mu, \qquad \xi = 1, \cdots, r\,;\, \eta = 1, \cdots, s$$

より

$$\mathscr{E}(\bar{x}..) = \mu$$

は明らか.

(8.6.9) の S の期待値を求めるには，どの S も次の4つの型の線形結合で表わされることに注意するとよい.

(8.6.10)
$$\begin{aligned} \mathscr{E}(x'_{\xi\eta}x'_{\eta'\xi'}) &= \sigma^2_{\cdot\cdot} + \sigma^2_{\cdot 0} + \sigma^2_{0\cdot}, & \xi = \xi', \eta = \eta' \\ &= \sigma^2_{\cdot 0}, & \xi = \xi', \eta \neq \eta' \\ &= \sigma^2_{0\cdot}, & \xi \neq \xi', \eta = \eta' \\ &= 0, & \xi \neq \xi', \eta \neq \eta' \end{aligned}$$

よって，期待値をとると

(8.6.11)
$$\begin{aligned} \mathscr{E}(S_{\cdot 0}) &= s(r-1)\left(\sigma^2_{\cdot 0} + \frac{\sigma^2_{\cdot\cdot}}{s}\right) \\ \mathscr{E}(S_{0\cdot}) &= r(s-1)\left(\sigma^2_{0\cdot} + \frac{\sigma^2_{\cdot\cdot}}{r}\right) \\ \mathscr{E}(S_{\cdot\cdot}) &= (r-1)(s-1)\sigma^2_{\cdot\cdot} \end{aligned}$$

一方，$\mathscr{E}(S_T)$ は

$$\mathscr{E}(S_T) = \mathscr{E}(S_{\cdot 0}) + \mathscr{E}(S_{\cdot 0}) + \mathscr{E}(S_{\cdot\cdot})$$

から求められる.

同様に，$\bar{x}..$ の分散も (8.6.10) の4つの型の線形結合で表わして，期待値をとると

(8.6.12)
$$\sigma^2(\bar{x}..) = \frac{\sigma^2_{\cdot 0}}{r} + \frac{\sigma^2_{0\cdot}}{s} + \frac{\sigma^2_{\cdot\cdot}}{rs}$$

が得られる．要約すると次を得る.

8.6.2　$(u_1, \cdots, u_r)$，$(v_1, \cdots, v_s)$ は無限母集団からの独立な標本とし，$(x_{\xi\eta}\,;\, \xi = 1, \cdots, r\,;\, \eta = 1, \cdots, s)$ を行列標本とする．(8.6.7) で $\bar{x}..$ を定義し，$S_{\cdot 0}$, $S_{0\cdot}$, $S_{\cdot\cdot}$ は (8.6.9) で与えられる全平方和 S_T の行要素，列要素，残差要素とする．このとき，$\mathscr{E}(\bar{x}..) = \mu$，$\sigma^2(\bar{x}..)$ は (8.6.12) で，$S_{\cdot 0}$, $S_{0\cdot}$, $S_{\cdot\cdot}$ の期待値は (8.6.11) で与えられる.

(b)　行と列を構成する母集団がともに有限な場合

$(u_1, \cdots, u_r)$，$(v_1, \cdots, v_s)$ がそれぞれ有限母集団 π_{N_1}，π_{N_2} からの独立な確率標本の場合は，前の議論が適用できる．この有限な場合は Hooke (1956 *a*)，Tukey (1950) によって論じられている.

有限母集団のときは，一般性を失うことなく，π_{N_1} の要素はすべて異なっている．すなわち，u の値は $u_{o1} < \cdots < u_{oN_1}$ と仮定してよい．同様に，v の値についても，$v_{o1} < \cdots < v_{oN_2}$ を仮定する．次の記号を用いよう.

(8.6.13)
$$x(u_{oi}, v_{oj}) = x_{oij}, \qquad x'_{oij} = x_{oij} - \mu_f$$

ただし，(8.6.1) に対応して

(8.6.14)
$$\mu_f = \frac{1}{N_1N_2}\sum_{i,j} x_{oij}, \quad \mu_{fi\cdot} = \frac{1}{N_2}\sum_{j} x_{oij}, \quad \mu_{f\cdot j} = \frac{1}{N_1}\sum_{i} x_{oij}$$

とする．有限な行と列からなる母集団に対しては，σ^2_f，$\sigma^2_{f\cdot 0}$，$\sigma^2_{f0\cdot}$，$\sigma^2_{f\cdot\cdot}$ を次で定義すると，$S_{\cdot 0}$，$S_{0\cdot}$，$S_{\cdot\cdot}$ の期待値を求め易くなるだろう.

$$\sigma^2_f = \frac{1}{N_1N_2 - 1}\sum_{i,j} x'^2_{oij}$$

(8.6.15)
$$\begin{aligned} \sigma^2_{f\cdot 0} &= \frac{1}{N_2(N_1 - 1)}\sum_{i,j} (\mu_{fi\cdot} - \mu_f)^2 \\ \sigma^2_{f0\cdot} &= \frac{1}{N_1(N_2 - 1)}\sum_{i,j} (\mu_{f\cdot j} - \mu_f)^2 \\ \sigma^2_{f\cdot\cdot} &= \frac{1}{(N_1 - 1)(N_2 - 1)}\sum_{i,j} (x_{oij} - \mu_{fi\cdot} - \mu_{f\cdot j} + \mu_f)^2 \end{aligned}$$

ただし σ^2_f，$\sigma^2_{f\cdot 0}$，$\sigma^2_{f0\cdot}$，$\sigma^2_{f\cdot\cdot}$ の間には，次の関係式が成立する.

(8.6.16)

$$\frac{N_1N_2 - 1}{N_1N_2}\sigma^2_f = \frac{(N_1 - 1)}{N_1}\sigma^2_{f\cdot 0} + \frac{(N_2 - 1)}{N_2}\sigma^2_{f0\cdot} + \frac{(N_1 - 1)(N_2 - 1)}{N_1N_2}\sigma^2_{f\cdot\cdot}$$

ここで $N_1, N_2 \longrightarrow \infty$ とすれば，(8.6.16) は (8.6.4 *a*) になることに注意せよ.

$\bar{x}..$ の期待値と分散および $S_{0\cdot}$，$S_{\cdot 0}$，$S_{\cdot\cdot}$ の期待値を求めよう．まず，次の量を定義しておくと便利である.

(8.6.17)
$$\begin{aligned} T_{(\cdot\cdot)} &= \sum_{i,j} (x'_{oij})^2, & T_{(\cdot 0)} &= \sum_{i,j\neq j'} x'_{oij}x'_{oij'} \\ T_{(0\cdot)} &= \sum_{i\neq i',j} x'_{oij}x'_{oi'j}, & T_{(00)} &= \sum_{i\neq i',j\neq j'} x'_{oij}x'_{oi'j'} \end{aligned}$$

このとき

(8.6.18)
$$T_{(\cdot\cdot)} + T_{(\cdot 0)} + T_{(0\cdot)} + T_{(00)} = \left[\sum_{i,j} x'_{oij}\right]^2 = 0$$

が成り立つ．

さて，$\sigma^2_{f\cdot 0}$，$\sigma^2_{f0\cdot}$，$\sigma^2_{f\cdot\cdot}$ は $T_{(\cdot\cdot)}$，$T_{(\cdot 0)}$，$T_{(0\cdot)}$，$T_{(00)}$ の線形関数だから，T について解き，(8.6.18) を用いると，次が得られる．

$$
\begin{aligned}
T_{(\cdot\cdot)} &= (N_1-1)(N_2-1)\left[\sigma^2_{f\cdot\cdot}+\frac{N_1}{N_1-1}\sigma^2_{f0\cdot}+\frac{N_2}{N_2-1}\sigma^2_{f\cdot 0}\right]\\
T_{(\cdot 0)} &= (N_1-1)(N_2-1)\left[-\sigma^2_{f\cdot\cdot}-\frac{N_1}{N_1-1}\sigma^2_{f0\cdot}+N_2\sigma^2_{f\cdot 0}\right]\\
T_{(0\cdot)} &= (N_1-1)(N_2-1)\left[-\sigma^2_{f\cdot\cdot}+N_1\sigma^2_{f0\cdot}-\frac{N_2}{N_2-1}\sigma^2_{f\cdot 0}\right]\\
T_{(00)} &= (N_1-1)(N_2-1)\left[\sigma^2_{f\cdot\cdot}-N_1\sigma^2_{f0\cdot}-N_2\sigma^2_{f\cdot 0}\right]
\end{aligned}
\tag{8.6.19}
$$

無限母集団と同様に，$(x_{\xi\eta};\xi=1,\cdots,r;\eta=1,\cdots,s)$ を行列標本とする．$T_{(\cdot\cdot)}$，$T_{(\cdot 0)}$，$T_{(0\cdot)}$，$T_{(00)}$ を x'_{oij} から定義したように，$T_{\cdot\cdot}$，$T_{\cdot 0}$，$T_{0\cdot}$，T_{00} を $x'_{\xi\eta}$ ((8.6.7) で与えられている) で定義する．しかし，この場合，$T_{\cdot\cdot}$，$T_{\cdot 0}$，$T_{0\cdot}$，T_{00} なる確率変数は，(8.6.18) を満足しないことに注意せよ．

容易に

$$
\begin{aligned}
\mathscr{E}(T_{\cdot\cdot}) &= \frac{rs}{N_1N_2}T_{(\cdot\cdot)}\\
\mathscr{E}(T_{\cdot 0}) &= \frac{rs(s-1)}{N_1N_2(N_2-1)}T_{(\cdot 0)}\\
\mathscr{E}(T_{0\cdot}) &= \frac{rs(r-1)}{N_1N_2(N_1-1)}T_{(0\cdot)}\\
\mathscr{E}(T_{00}) &= \frac{rs(r-1)(s-1)}{N_1N_2(N_1-1)(N_2-1)}T_{(00)}
\end{aligned}
\tag{8.6.20}
$$

が確かめられる．(8.6.20) は，すぐあとで見るように，$S_{\cdot 0}$，$S_{0\cdot}$，$S_{\cdot\cdot}$ の期待値を求める問題には欠かせない式である．

さて，$S_{\cdot 0}$，$S_{0\cdot}$，$S_{\cdot\cdot}$ は，次のような T_{00}，$T_{\cdot 0}$，$T_{0\cdot}$，$T_{\cdot\cdot}$ の1次式になっている．

$$
\begin{aligned}
S_{\cdot 0} &= \frac{r-1}{rs}(T_{\cdot\cdot}+T_{\cdot 0})-\frac{1}{rs}(T_{0\cdot}+T_{00})\\
S_{0\cdot} &= \frac{s-1}{rs}(T_{\cdot\cdot}+T_{0\cdot})-\frac{1}{rs}(T_{\cdot 0}+T_{00})\\
S_{\cdot\cdot} &= \frac{(r-1)(s-1)}{rs}T_{\cdot\cdot}-\frac{(r-1)}{rs}T_{\cdot 0}-\frac{(s-1)}{rs}T_{0\cdot}+\frac{1}{rs}T_{00}
\end{aligned}
\tag{8.6.21}
$$

期待値をとり，(8.6.20) を用いる．$T_{(\cdot\cdot)}$，$T_{(\cdot 0)}$，$T_{(00)}$ に (8.6.19) を代入すると，S の期待値

$$
\begin{aligned}
\mathscr{E}(S_{\cdot 0}) &= s(r-1)\left[\sigma^2_{f\cdot 0}+\left(\frac{1}{s}-\frac{1}{N_2}\right)\sigma^2_{f\cdot\cdot}\right]\\
\mathscr{E}(S_{0\cdot}) &= r(s-1)\left[\sigma^2_{f0\cdot}+\left(\frac{1}{r}-\frac{1}{N_1}\right)\sigma^2_{f\cdot\cdot}\right]\\
\mathscr{E}(S_{\cdot\cdot}) &= (r-1)(s-1)\sigma^2_{f\cdot\cdot}
\end{aligned}
\tag{8.6.22}
$$

が得られる．$\mathscr{E}(S_T)$ の値は

$$\mathscr{E}(S_T)=\mathscr{E}(S_{\cdot 0})+\mathscr{E}(S_{0\cdot})+\mathscr{E}(S_{\cdot\cdot})$$

になる．$N_1, N_2\longrightarrow\infty$ としたとき，$\sigma^2_{f\cdot\cdot}$，$\sigma^2_{f0\cdot}$，$\sigma^2_{f\cdot 0}$ がそれぞれ極値として $\sigma^2_{\cdot\cdot}$，$\sigma^2_{0\cdot}$，$\sigma^2_{\cdot 0}$ を持つならば，上式は (8.6.11) に帰着する．

$\bar{x}_{\cdot\cdot}$ の分散は $(\bar{x}_{\cdot\cdot}-\mu_f)^2$ の期待値である．ところが $(\bar{x}_{\cdot\cdot}-\mu_f)^2$ は

$$(\bar{x}_{\cdot\cdot}-\mu_f)^2=\frac{1}{rs}[T_{\cdot\cdot}+T_{\cdot 0}+T_{0\cdot}+T_{00}] \tag{8.6.23}$$

と表わされる．この両辺の期待値をとって，(8.6.19)，(8.6.20) を用いれば

(8.6.24)

$$\sigma^2(\bar{x}_{\cdot\cdot})=\left(\frac{1}{r}-\frac{1}{N_1}\right)\sigma^2_{f\cdot 0}+\left(\frac{1}{s}-\frac{1}{N_2}\right)\sigma^2_{f0\cdot}+\left(\frac{1}{r}-\frac{1}{N_1}\right)\left(\frac{1}{s}-\frac{1}{N_2}\right)\sigma^2_{f\cdot\cdot}$$

を得る．読者は，$N_1, N_2\longrightarrow\infty$ のとき，(8.6.24) が (8.6.12) になることに注意せよ．また，有限な行からなる母集団および有限な列からなる母集団の場合について，**8.6.2** のように，結果を要約してみよ．

(c)　3次の行列標本

(a)，(b) での考え方は3次あるいはより高次の行列標本の場合にも，比較的容易に拡張できる．高次のときは，分散が多種類になるので，本質的には，3次の場合を考えれば十分であろう．ここでは，u，v，w が c.d.f. $F_1(u)$，$F_2(v)$，$F_3(w)$ を持つ独立な確率変数として，確率変数 $x(u,v,w)$ を考察しよう．

$x(u,v,w)$ に対して，次の期待値と確率変数を定義する．

$$\mu=\mathscr{E}(x(u,v,w))$$

$$
\begin{aligned}
\mu_{u\cdot\cdot} &= \int_{R_2}x(u,v,w)\,dF_2(v)\,dF_3(w), &&\mu_{\cdot v\cdot},\ \mu_{\cdot\cdot w}\ \text{に対しても同様}\\
\mu_{uv\cdot} &= \int_{-\infty}^{\infty}x(u,v,w)\,dF_3(w), &&\mu_{u\cdot w},\ \mu_{\cdot vw}\ \text{に対しても同様}
\end{aligned}
\tag{8.6.25}
$$

$$\varepsilon_{u\cdot\cdot} = \mu_{u\cdot\cdot} - \mu, \qquad \varepsilon_{\cdot v\cdot},\ \varepsilon_{\cdot\cdot w}\ \text{に対しても同様}$$

(8.6.26) $$\varepsilon_{uv\cdot} = \mu_{uv\cdot} - \mu_{u\cdot\cdot} - \mu_{\cdot v\cdot} + \mu, \qquad \varepsilon_{u\cdot w},\ \varepsilon_{\cdot vw}\ \text{に対しても同様}$$

$$\varepsilon_{uvw} = x(u,v,w) - \mu_{uv\cdot} - \mu_{u\cdot w} - \mu_{\cdot vw} + \mu_{u\cdot\cdot} + \mu_{\cdot v\cdot} + \mu_{\cdot\cdot w} - \mu$$

8.6.1 の拡張として，次の **8.6.3** が成立する．

8.6.3　u, v, w が独立な確率変数で，かつ $x(u,v,w)$ が確率変数であれば，

$$x(u,v,w) = \mu + \varepsilon_{u\cdot\cdot} + \varepsilon_{\cdot v\cdot} + \varepsilon_{\cdot\cdot w} + \varepsilon_{uv\cdot} + \varepsilon_{u\cdot w} + \varepsilon_{\cdot vw} + \varepsilon_{uvw}$$

が成り立つ．ただし，各 ε は (8.6.26) で定義され，期待値 0，共分散 0 を持ち

(8.6.27) $$\sigma^2(x(u,v,w)) = \sigma^2(\varepsilon_{u\cdot\cdot}) + \sigma^2(\varepsilon_{\cdot v\cdot}) + \sigma^2(\varepsilon_{\cdot\cdot w}) + \sigma^2(\varepsilon_{uv\cdot}) + \sigma^2(\varepsilon_{u\cdot w}) + \sigma^2(\varepsilon_{\cdot vw}) + \sigma^2(\varepsilon_{uvw})$$

を満たす．

便宜上，次の記号を用いれば，(8.6.27) は

$$\sigma^2(x(u,v,w)) = \sigma^2$$

$$\sigma^2(\varepsilon_{u\cdot\cdot}) = \sigma^2_{\cdot 00}, \qquad \sigma^2(\varepsilon_{\cdot v\cdot}) = \sigma^2_{0\cdot 0}, \qquad \sigma^2(\varepsilon_{\cdot\cdot w}) = \sigma^2_{00\cdot}$$

$$\sigma^2(\varepsilon_{uv\cdot}) = \sigma^2_{\cdot\cdot 0}, \qquad \sigma^2(\varepsilon_{u\cdot w}) = \sigma^2_{\cdot 0\cdot}, \qquad \sigma^2(\varepsilon_{\cdot vw}) = \sigma^2_{0\cdot\cdot}$$

$$\sigma^2(\varepsilon_{uvw}) = \sigma^2_{\cdot\cdot\cdot}$$

(8.6.28) $$\sigma^2 = \sigma^2_{\cdot 00} + \sigma^2_{0\cdot 0} + \sigma^2_{00\cdot} + \sigma^2_{\cdot\cdot 0} + \sigma^2_{\cdot 0\cdot} + \sigma^2_{0\cdot\cdot} + \sigma^2_{\cdot\cdot\cdot}$$

と書き表わされる．

いま，$(u_1, \cdots, u_r)$，$(v_1, \cdots, v_s)$，$(w_1, \cdots, w_t)$ をそれぞれ c.d.f. $F_1(u)$，$F_2(v)$，$F_3(w)$ を持つ無限母集団からの独立な標本としよう．この3つを合わせた標本の c.d.f. は

(8.6.29) $$\prod_{\xi=1}^{r} F_1(u_\xi) \cdot \prod_{\eta=1}^{s} F_2(v_\eta) \cdot \prod_{\zeta=1}^{t} F_3(w_\zeta)$$

である．3次の行列標本は確率変数の組

$$(x(u_\xi, v_\eta, w_\zeta)\ ; \xi = 1, \cdots, r\ ;\quad \eta = 1, \cdots, s\ ;\quad \zeta = 1, \cdots, t)$$

からなり，r 個の行，s 個の列，t 個の層を持つ3次元の行列とみなすことができる．

この標本に対して，定義 (8.6.7) を拡張して，$x_{\xi\eta\zeta}$，$x'_{\xi\eta\zeta}$，および標本平均 $\bar{x}_{\cdot\cdot\cdot}$，$\bar{x}_{\xi\cdot\cdot}$，$\bar{x}_{\cdot\eta\cdot}$，$\bar{x}_{\cdot\cdot\zeta}$，$\bar{x}_{\xi\eta\cdot}$，$\bar{x}_{\xi\cdot\zeta}$，$\bar{x}_{\cdot\eta\zeta}$ を定義する．各確率変数 $x_{\xi\eta\zeta}$ は **8.6.3** で述べられた性質を持つ確率変数の和であるから

(8.6.30) $$x_{\xi\eta\zeta} = \mu + \varepsilon_{\xi\cdot\cdot} + \varepsilon_{\cdot\eta\cdot} + \varepsilon_{\cdot\cdot\zeta} + \varepsilon_{\xi\eta\cdot} + \varepsilon_{\xi\cdot\zeta} + \varepsilon_{\cdot\eta\zeta} + \varepsilon_{\xi\eta\zeta}$$

と表わせる．

さらに，平方和

(8.6.31)

$$S_T = \sum_{\xi,\eta,\zeta} (x_{\xi\eta\zeta} - \bar{x}_{\cdot\cdot\cdot})^2$$

$$S_{\cdot 00} = \sum_{\xi,\eta,\zeta} (\bar{x}_{\xi\cdot\cdot} - \bar{x}_{\cdot\cdot\cdot})^2, \qquad S_{0\cdot 0},\ S_{00\cdot}\ \text{に対しても同様}$$

$$S_{\cdot\cdot 0} = \sum_{\xi,\eta,\zeta} (\bar{x}_{\xi\eta\cdot} - \bar{x}_{\xi\cdot\cdot} - \bar{x}_{\cdot\eta\cdot} + \bar{x}_{\cdot\cdot\cdot})^2, \qquad S_{\cdot 0\cdot},\ S_{0\cdot\cdot}\ \text{に対しても同様}$$

$$S_{\cdot\cdot\cdot} = \sum_{\xi,\eta,\zeta} (x_{\xi\eta\zeta} - \bar{x}_{\xi\eta\cdot} - \bar{x}_{\xi\cdot\zeta} - \bar{x}_{\cdot\eta\zeta} + \bar{x}_{\xi\cdot\cdot} + \bar{x}_{\cdot\eta\cdot} + \bar{x}_{\cdot\cdot\zeta} - \bar{x}_{\cdot\cdot\cdot})^2$$

を定義すれば

(8.6.32) $$S_T = S_{\cdot 00} + S_{0\cdot 0} + S_{00\cdot} + S_{\cdot\cdot 0} + S_{\cdot 0\cdot} + S_{0\cdot\cdot} + S_{\cdot\cdot\cdot}$$

が確かめられる．

ここでも，$S_{\cdot 00}$，$S_{0\cdot 0}$，$S_{00\cdot}$ をそれぞれ S_T の行要素，列要素，層要素といい，$S_{\cdot\cdot 0}$，$S_{\cdot 0\cdot}$，$S_{0\cdot\cdot}$ を S_T の行-列間相互作用要素，列-層間相互作用要素，層-行間相互作用要素，$S_{\cdot\cdot\cdot}$ を S_T の 残差要素と呼ぶ．

明らかに

$$\mathscr{E}(\bar{x}_{\cdot\cdot\cdot}) = \mu$$

各 S の期待値および $\bar{x}_{\cdot\cdot\cdot}$ の分散を求めるために，これらの量はいずれも，次の量の線形結合で表わせることを注意しておこう．

(8.6.33)
$$\begin{aligned}
\mathscr{E}(x'_{\xi\eta\zeta}x'_{\xi'\eta'\zeta'}) &= \sigma^2, & \xi &= \xi', & \eta &= \eta', & \zeta &= \zeta' \\
&= \sigma^2_{\cdot\cdot 0} + \sigma^2_{\cdot 00} + \sigma^2_{0\cdot 0}, & \xi &= \xi', & \eta &= \eta', & \zeta &\fallingdotseq \zeta' \\
&= \sigma^2_{\cdot 0\cdot} + \sigma^2_{\cdot 00} + \sigma^2_{00\cdot}, & \xi &= \xi', & \eta &\fallingdotseq \eta', & \zeta &= \zeta' \\
&= \sigma^2_{0\cdot\cdot} + \sigma^2_{0\cdot 0} + \sigma^2_{00\cdot}, & \xi &\fallingdotseq \xi', & \eta &= \eta', & \zeta &= \zeta' \\
&= \sigma^2_{\cdot 00}, & \xi &= \xi', & \eta &\fallingdotseq \eta', & \zeta &\fallingdotseq \zeta' \\
&= \sigma^2_{0\cdot 0}, & \xi &\fallingdotseq \xi', & \eta &= \eta', & \zeta &\fallingdotseq \zeta' \\
&= \sigma^2_{00\cdot}, & \xi &\fallingdotseq \xi', & \eta &\fallingdotseq \eta', & \zeta &= \zeta' \\
&= 0, & \xi &\fallingdotseq \xi', & \eta &\fallingdotseq \eta', & \zeta &\fallingdotseq \zeta'
\end{aligned}$$

ただし σ^2 は (8.6.28) で与えられている．

S の期待値を評価すると

$$\mathscr{E}(S_{00\cdot}) = rs(t-1)\left[\sigma^2_{00\cdot} + \frac{1}{s}\sigma^2_{0\cdot\cdot} + \frac{1}{r}\sigma^2_{\cdot 0\cdot} + \frac{1}{rs}\sigma^2_{\cdot\cdot\cdot}\right]$$

$$\mathscr{E}(S_{0\cdot 0}) = rt(s-1)\left[\sigma^2_{0\cdot 0} + \frac{1}{t}\sigma^2_{0\cdot\cdot} + \frac{1}{r}\sigma^2_{\cdot\cdot 0} + \frac{1}{rt}\sigma^2_{\cdot\cdot\cdot}\right]$$

$$\mathscr{E}(S_{\cdot 00}) = st(r-1)\left[\sigma^2_{\cdot 00} + \frac{1}{t}\sigma^2_{\cdot 0\cdot} + \frac{1}{s}\sigma^2_{\cdot\cdot 0} + \frac{1}{st}\sigma^2_{\cdot\cdot\cdot}\right]$$

(8.6.34)
$$\mathscr{E}(S_{0\cdot\cdot}) = r(s-1)(t-1)\left[\sigma^2_{0\cdot\cdot} + \frac{1}{r}\sigma^2_{\cdot\cdot\cdot}\right]$$

$$\mathscr{E}(S_{\cdot 0\cdot}) = s(r-1)(t-1)\left[\sigma^2_{\cdot 0\cdot} + \frac{1}{s}\sigma^2_{\cdot\cdot\cdot}\right]$$

$$\mathscr{E}(S_{\cdot\cdot 0}) = t(r-1)(s-1)\left[\sigma^2_{\cdot\cdot 0} + \frac{1}{t}\sigma^2_{\cdot\cdot\cdot}\right]$$

$$\mathscr{E}(S_{\cdot\cdot\cdot}) = (r-1)(s-1)(t-1)\sigma^2_{\cdot\cdot\cdot}$$

もちろん，$\mathscr{E}(S_T)$ は (8.6.34) の右辺の和である．

$\bar{x}_{\cdots}$ の分散も，同様に (8.6.33) の量の線形結合で表わされるから

(8.6.35)
$$\sigma^2(\bar{x}_{\cdots}) = \frac{1}{r}\sigma^2_{\cdot 00} + \frac{1}{s}\sigma^2_{0\cdot 0} + \frac{1}{t}\sigma^2_{00\cdot} + \frac{1}{rs}\sigma^2_{\cdot\cdot 0} + \frac{1}{rt}\sigma^2_{\cdot 0\cdot} + \frac{1}{st}\sigma^2_{0\cdot\cdot} + \frac{1}{rst}\sigma^2_{\cdot\cdot\cdot}$$

が求まる．

読者は **8.6.2** を拡張して，それぞれ無限の行，列，層からなる母集団に対して，3次の行列標本の場合について公式化せよ．

(8.6.11) 式と (8.6.22) 式の構造を比較すれば，$(u_1, \cdots, u_r)$，$(v_1, \cdots, v_s)$，$(w_1, \cdots, w_t)$ がそれぞれ有限な行，列，層からなる母集団からの独立な標本のときに，行列標本 $\{x(u_\xi, v_\eta, w_\zeta)\}$ に対して (8.6.34) 式がどのような型になるか容易に推測できるであろう．同様に，(8.6.12) 式と (8.6.24) 式を比べると，$\sigma^2(\bar{x}_{\cdots})$ に対する (8.6.35) 式が，有限母集団の場合には，どのような型になるかもわかる．これを試みて，公式にまとめてみよ．

(d)　釣合い型不完備行列標本

2次の行列標本 $(x_{\xi\eta}\,;\xi = 1, \cdots, r\,;\eta = 1, \cdots, s)$ に戻って考えよう．この標本を $\{x_{\xi\eta}\}$ と簡単に表わす．この行から s' 個の要素，列から r' 個の要素を選び，部分標本 $\{x_{\xi\eta}\}^*$ を取り出す．ただし，n を標本の大きさとすれば，n，r，s，r'，s' は条件

(8.6.36)
$$n = rs' = r's$$

を満たす正の整数でなければならない．これを満たす (ξ, η) の集合を G^* で表わす．この標本を釣合い型不完備行列標本と呼ぶ．G^* がつくられる例としては；$n = 12$，$r = 3$，$s = 6$，$r' = 2$，$s' = 4$；$n = 16$，$r = 4$，$s = 8$，$r' = 2$，$s' = 4$；$n = 40$，$r = 5$，$s = 10$，$r' = 4$，$s' = 8$ などがあげられる．部分標本の選び方については，Bose (1939)，Connor (1952)，Shrikhande (1952)，Cochran と Cox (1957)，Kempthorne (1952) で詳しく論じられている．

不完備標本の平均，ξ 番目の行平均，η 番目の列平均をそれぞれ $\bar{x}^*_{\cdot\cdot}$，$\bar{x}^*_{\xi\cdot}$，$\bar{x}^*_{\cdot\eta}$ で表わし

(8.6.37)
$$S^*_T = \sum_{\xi,\eta}{}^*(x_{\xi\eta} - \bar{x}^*_{\cdot\cdot})^2, \qquad S^*_{\cdot 0} = \sum_{\xi,\eta}{}^*(\bar{x}^*_{\xi\cdot} - \bar{x}^*_{\cdot\cdot})^2$$
$$S^*_{0\cdot} = \sum_{\xi,\eta}{}^*(x^*_{\cdot\eta} - \bar{x}^*_{\cdot\cdot})^2, \qquad S^*_E = S^*_T - S^*_{\cdot 0} - S^*_{0\cdot}$$

とする．ただし $\sum_{\xi,\eta}{}^*$ は $(\xi, \eta) \in G^*$ にわたる和を示す．S^*_E は残差平方和である．

行および列が無限の要素からなる母集団について，$\bar{x}^*_{\cdot\cdot}$ の分散，S^* の期待値を求めよう．ここでも，各期待値は (8.6.10) と同様な量の線形結合である．$\sigma^2(\bar{x}^*_{\cdot\cdot})$，$\mathscr{E}(S^*_T)$，$\mathscr{E}(S^*_{\cdot 0})$，$\mathscr{E}(S^*_{0\cdot})$ も容易に求まり

(8.6.38)
$$\mathscr{E}(S^*_E) = \mathscr{E}(S^*_T) - \mathscr{E}(S^*_{\cdot 0}) - \mathscr{E}(S^*_{0\cdot})$$

である．

$rs' = r's = n$ とすれば

(8.6.39)
$$\mathscr{E}(\bar{x}^*_{\cdot\cdot}) = \mu$$
$$\sigma^2(\bar{x}^*_{\cdot\cdot}) = \frac{1}{n}\sigma^2_{\cdot\cdot} + \frac{1}{r}\sigma^2_{\cdot 0} + \frac{1}{s}\sigma^2_{0\cdot}$$
$$\mathscr{E}(S^*_{\cdot 0}) = (r-1)\sigma^2_{\cdot\cdot} + (r-1)s'\sigma^2_{\cdot 0} + (r-r')\sigma^2_{0\cdot}$$
$$\mathscr{E}(S^*_{0\cdot}) = (s-1)\sigma^2_{\cdot\cdot} + (s-s')\sigma^2_{\cdot 0} + (s-1)r'\sigma^2_{0\cdot}$$
$$\mathscr{E}(S^*_E) = (n-r-s+1)\sigma^2_{\cdot\cdot} - (s-s')\sigma^2_{\cdot 0} - (r-r')\sigma^2_{0\cdot}$$

になる．

$r' = r$，$s' = s$ のとき，(8.6.39) は (8.6.12) と (8.6.11) に帰着する．

釣合い型不完備行列標本は，3次および高次の行列標本からの一般型において，幾種類もの条件を満たさなければならないので，結果的には，かなり複雑になってくる．しかし，ここでは，実験計画法において特に基本的な3次の場合だけを，しかも，特殊な場合だけに限定して考察してみよう．

r 個の行，r 個の列，r 個の層を持つ3次の行列標本 $\{x_{\xi\eta\zeta}\}$，$\xi,\eta,\zeta=1,\cdots,r$ を考える．釣合い型不完備標本 $\{x_{\xi\eta\zeta}\}^*$ として，(ξ,η) の各組に対してただ1つの要素，(η,ζ) の各組からただ1つの要素，そして (ζ,ξ) に対しても1つの要素からなる $\{x_{\xi\eta\zeta}\}$ の部分集合を選ぶ．

$\{x_{\xi\eta\zeta}\}^*$ は r^2 個の $\{x_{\xi\eta\zeta}\}$ の要素よりなっている．しかも $\{x_{\xi\eta\zeta}\}^*$ の要素は完全にバランスがとれている．事実，$\{x_{\xi\eta\zeta}\}^*$ のすべての要素を $\xi\eta$ 平面（行-列平面）上へ射影すれば，$\{x_{\xi\eta\zeta}\}^*$ の射影された各行には，$\{x_{\xi\eta\zeta}\}$ の層のおのおのからの要素がただ1つある．同じく，$\xi\eta$ 平面の各列にも，$\{x_{\xi\eta\zeta}\}$ の各層からの要素をただ1つ見い出すことができる．さらに，$\{x_{\xi\eta\zeta}\}^*$ を $\xi\zeta$ 平面，および $\eta\zeta$ 平面上へ射影したときも，“行”，“列”，“層”を適当に入れ替えれば，同様な事柄が成り立つ．$\{x_{\xi\eta\zeta}\}^*$ の射影された型を行，列および層からの**ラテン方格選択**と呼んでいる．

さて，不完備標本からのこのような特殊な選択による (ξ,η,ζ) の集合を G^* で表わそう．いうまでもなく，G^* の要素の選び方は多い．集合 G^* の構成（ラテン方格の構成）に興味ある読者は Bose (1938)，Mann (1943) を参照せよ．

ここでは，不完備標本における標本平均 $\bar{x}^*_{\cdot\cdot\cdot}$ および ξ 番目の行平均 $\bar{x}^*_{\xi\cdot\cdot}$，$\eta$ 番目の列平均 $\bar{x}^*_{\cdot\eta\cdot}$，$\zeta$ 番目の層平均 $\bar{x}^*_{\cdot\cdot\zeta}$ を考えよう．さらに，次の量を定義しておく．

(8.6.40)
$$
\begin{aligned}
S^*_T &= \sum_{\xi,\eta,\zeta}{}^*(x_{\xi\eta\zeta}-\bar{x}^*_{\cdot\cdot\cdot})^2, \qquad S^*_{\cdot 00} = \sum_{\xi,\eta,\zeta}{}^*(\bar{x}^*_{\xi\cdot\cdot}-\bar{x}^*_{\cdot\cdot\cdot})^2\\
S^*_{0\cdot 0} &= \sum_{\xi,\eta,\zeta}{}^*(\bar{x}^*_{\cdot\eta\cdot}-\bar{x}^*_{\cdot\cdot\cdot})^2, \qquad S^*_{00\cdot} = \sum_{\xi,\eta,\zeta}{}^*(\bar{x}^*_{\cdot\cdot\zeta}-\bar{x}^*_{\cdot\cdot\cdot})^2\\
S^*_E &= S^*_T - S^*_{\cdot 00} - S^*_{0\cdot 0} - S^*_{00\cdot}.
\end{aligned}
$$

ただし $\sum_{\xi,\eta,\zeta}{}^*$ は $(\xi,\eta,\zeta)\in G^*$ なる (ξ,η,ζ) のすべての和を示す．

$\bar{x}^*_{\cdot\cdot\cdot}$ の分散と (8.6.40) の各 S^* の期待値については，(8.6.33) の最初の4つの量だけの線形結合で表わされることに注目して，期待値をとると

(8.6.41)
$$
\begin{aligned}
\mathscr{E}(S^*_{\cdot 00}) &= (r-1)\sigma^2_E + (r-1)\sigma^2_{\cdot 00}\\
\mathscr{E}(S^*_{0\cdot 0}) &= (r-1)\sigma^2_E + (r-1)\sigma^2_{0\cdot 0}\\
\mathscr{E}(S^*_{00\cdot}) &= (r-1)\sigma^2_E + (r-1)\sigma^2_{00\cdot}\\
\mathscr{E}(S^*_E) &= (r-1)(r-2)\sigma^2_E + (r-1)^2[\sigma^2_{\cdot 00}+\sigma^2_{0\cdot 0}+\sigma^2_{00\cdot}]
\end{aligned}
$$

ただし σ^2_E は

(8.6.42)
$$\sigma^2_{\cdot} = \sigma^2_E + \sigma^2_{\cdot 00} + \sigma^2_{0\cdot 0} + \sigma^2_{00\cdot}$$

を満たす．(8.6.27)，(8.6.28) から σ^2_E は σ^2 の4つの和，すなわち

(8.6.43)
$$\sigma^2_E = \sigma^2_{\cdot\cdot\cdot} + \sigma^2_{\cdot\cdot 0} + \sigma^2_{\cdot 0\cdot} + \sigma^2_{0\cdot\cdot}$$

になることがわかる．

8.7　順序統計量に関する標本論

$(x_1,\cdots,x_n)$ は連続な c. d. f. $F(x)$ を持つ母集団からの標本と仮定する．$x_1,\cdots,x_n$ を小さい方から大きい方へと，左から右辺へ並べ換え，その順序づけられた値を $x_{(1)},\cdots,x_{(n)}$ とする．つまり $x_{(1)}\leqslant\cdots\leqslant x_{(n)}$ である．この新しい確率変数を**標本順序統計量**，$x_{(k)}$ を **k 番目の順序統計量** と呼ぶ．$F(x)$ は連続だから，$P(x_{(\xi-1)}=x_{(\xi)})=0$，$\xi=1,\cdots,n$ であることに注意しよう．区間 $(-\infty, x_{(1)}]$，$(x_{(1)}, x_{(2)}]$，$\cdots$，$(x_{(n)},+\infty)$ をそれぞれ**標本ブロック** $B_1^{(1)},\cdots,B_1^{(n+1)}$ と呼び，このブロックの関数 $F(x_{(1)})$，$F(x_{(2)})-F(x_{(1)}),\cdots,1-F(x_{(n)})$ をそれぞれ**カバー**[*] $u_1,\cdots,u_{n+1}$ という．u の和は1だから，しばしば，u_{n+1} を省くことがある．B の右下の添字は次元を示している．2次元あるいは多次元の標本ブロックについては，8.7(c) 節で定義されよう．与えられた標本ブロックに対するカバーは，その標本ブロックの母集団分布（P 測度）による確率に他ならない．もちろん，$u_1,\cdots,u_{n+1}$ は確率変数である．

順序統計量およびカバーに関する標本論は，ノンパラメトリック統計的推定において特に重要である．これは第11，14章で詳しく論じる．この節では，順序統計量に関する基本的な結果をいくつか考察しよう．この方面に興味ある読者は Fraser (1957)，Gumbel (1958)，Kendall (1953) を参照せよ．特に Savage (1962) には，順序統計量に関する多くの文献が網羅されている．

(a) 順序統計量の標本分布

7.1.1 で述べたように，確率変数 x が連続な c. d. f. $F(x)$ を持てば，$y=F(x)$ なる確率変数 y は矩形分布 $R\left(\frac{1}{2},1\right)$ に従う．$y_\xi=F(x_\xi)$，$\xi=1,\cdots,n$ とすれば，$(y_1,\cdots,y_n)$ は矩形分布 $R\left(\frac{1}{2},1\right)$ からの大きさ n の標本となる．この y の p. e. は標本空間

[*] 原書では coverage である．日本語訳として，区間変量ともいわれているが，ここでは，以後，カバーと訳すことにする．（訳注）

R_n の単位立方体 $\{(y_1, \cdots, y_n) : 0 \leqslant y_\xi \leqslant 1,\ \xi = 1, \cdots, n\}$ 上で

(8.7.1) $$1 \cdot dy_1 \cdots dy_n.$$

その他では，$0 \cdot dy_1 \cdots dy_n$ である．$y_{(1)} \leqslant \cdots \leqslant y_{(n)}$ を標本 $(y_1, \cdots, y_n)$ からの順序統計量として，この順序統計量の p.e. を求めよう．標本 $(y_1, \cdots, y_n)$ の2つあるいはそれ以上の要素が等しくなる確率は0である．したがって，単位立方体の座標が異なる点だけ考えれば十分である．$(y_1, \cdots, y_n)$ をこのような点 P とする．点 P に関する p.e. は $1 \cdot dy_1 \cdots dy_n$ である．しかし，P の座標を置換することによって，全体で $n!$ 通りの単位立方体の点が得られる．この $n!$ 個の点のうち，1番小さい座標を $y_{(1)}$，2番目に小さい座標を $y_{(2)}, \cdots$，とすると，新しい確率変数が得られる．標本 $(y_1, \cdots, y_n)$ の順序統計量を要素に持つ確率変数 $(y_{(1)}, \cdots, y_{(n)})$ の p.d.f. は，この $n!$ 個の点の p.d.f. を加え合わせ，かつ順序統計量の記号を用いて表わされる．すなわち，$(y_{(1)}, \cdots, y_{(n)})$ の p.e. は領域 $\{y_{(1)}, \cdots, y_{(n)}) : 0 \leqslant y_{(1)} \leqslant \cdots \leqslant y_{(n)} \leqslant 1\}$ 内で

(8.7.2) $$n!\, dy_{(1)} \cdots dy_{(n)}$$

その他で，$0 \cdot dy_{(1)} \cdots dy_{(n)}$ である．これは (7.7.17) で定義された，順序づけられた n 変数ディリクレ分布 $D^*(1, \cdots, 1; 1)$ である．したがって，次の結果が成り立つ．

8.7.1　$(x_{(1)}, \cdots, x_{(n)})$ が連続な c.d.f. $F(x)$ からの標本順序統計量とすれば，確率変数 $F(x_{(1)}), \cdots, F(x_{(n)})$ は順序づけられた n 変数ディリクレ分布 $D^*(1, \cdots, 1; 1)$ を持つ．

確率変数 $F(x_{(1)}), \cdots, F(x_{(n)})$ のうち，任意の1つあるいはそれ以上の部分集合の分布は，**7.7.6** を適用すれば，すぐわかる．1つの場合には，次を得る．

8.7.2　確率変数 $F(x_{(k)})$，$1 \leqslant k \leqslant n$ はベータ分布 $Be(k, n-k+1)$ を持つ．

2つ以上の任意の $F(x_{(k)})$ の分布については，**8.7.3** が成立する．

8.7.3　確率変数 $F(x_{(k_1)}), F(x_{(k_1+k_2)}), \cdots, F(x_{(k_1+\cdots+k_s)})$ は順序づけられた s 変数ディリクレ分布 $D^*(k_1, \cdots, k_s; n-k_1-\cdots-k_s+1)$ に従う．

8.7.1，**8.7.2**，**8.7.3** においては，$F(x)$ が連続であると仮定していることに注意せよ．さらに，p.d.f. $f(x)$ が存在するという強い仮定の下では，順序統計量 $x_{(1)}, \cdots, x_{(n)}$

およびこの部分集合の p.e. を求めることができる．これは **8.7.1**, **8.7.2**, **8.7.3** に対応して，次のようになる．

8.7.1 a　$(x_{(1)}, \cdots, x_{(n)})$ が c.d.f. $F(x)$，p.d.f. $f(x)$ を持つ母集団からの大きさ n の標本順序統計量であるとき，この順序統計量の p.e. は $(x_{(1)}, \cdots, x_{(n)})$ の標本空間 R_n の $-\infty \leqslant x_{(1)} \leqslant \cdots \leqslant x_{(n)} \leqslant +\infty$ に対して

(8.7.3) $$n!\, f(x_{(1)}) \cdots f(x_{(n)})\, dx_{(1)} \cdots dx_{(n)}$$

になる．

8.7.2 a　**8.7.1 a** の仮定の下に，$x_{(k)}$，$1 \leqslant k \leqslant n$ の p.e. は

(8.7.4) $$\frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} [F(x_{(k)})]^{k-1} [1 - F(x_{(k)})]^{n-k} f(x_{(k)}) dx_{(k)}$$

で，$x_{(k)}$ の標本空間は R_1 である．

8.7.3 a　**8.7.1 a** の仮定の下で，$(x_{(k_1)}, x_{(k_1+k_2)}, \cdots, x_{(k_1+\cdots+k_s)})$ の p.e. は R_s の $-\infty \leqslant x_{(k_1)} \leqslant \cdots \leqslant x_{(k_1+\cdots+k_s)} \leqslant +\infty$ なる標本空間 $(x_{(k_1)}, x_{(k_1+k_2)}, \cdots, x_{(k_1+\cdots+k_s)})$ では

(8.7.5) $$\frac{\Gamma(n+1)}{\Gamma(k_1) \cdots \Gamma(k_s) \Gamma(n-k_1-\cdots-k_s+1)}$$
$$\cdot [F(x_{(k_1)})]^{k_1-1} [F(x_{(k_1+k_2)}) - F(x_{(k_1)})]^{k_2-1} \cdots$$
$$\cdot [1 - F(x_{(k_1+\cdots+k_s)})]^{n-k_1-\cdots-k_s} f(x_{(k_1)}) \cdots$$
$$\cdot f(x_{(k_1+\cdots+k_s)})\, dx_{(k_1)} \cdots dx_{(k_1+\cdots+k_s)}.$$

公式 (8.7.3)，(8.7.4)，(8.7.5) は Craig (1932) によってはじめて得られた．そこでは興味ある特殊な場合が多く議論されている．たとえば，(8.7.4) において，$k=1$ とすれば，標本最小要素の p.e. が得られる．$k=n$ のとき，標本最大要素の p.e. が見い出せる．n が奇数のとき，$k = \dfrac{n+1}{2}$ とおけば，標本メディアンの p.e. が求まる．(8.7.5) において，$s=2$，$k_1=1$，$k_2=n-1$ とすれば，標本最大要素と最小要素の p.e. は

(8.7.6) $$n(n-1)[F(x_{(n)}) - F(x_{(1)})]^{n-2} f(x_{(1)}) f(x_{(n)})\, dx_{(1)}\, dx_{(n)}.$$

標本範囲 w の分布は，変換

$$x_{(1)} = v, \qquad x_{(n)} = v + w$$

を (8.7.6) に施し，w の周辺分布を求めればよい．すなわち，変数 v で積分すれば，w

の分布が得られる．

(b) 1次元カバーの標本分布

8.7.1, 8.7.2, 8.7.3 における分布については，幅広い解釈が成り立つ．まず，次の変換を行なおう．

(8.7.7)　$y_{(1)} = u_1,\quad y_{(2)} = u_1 + u_2,\quad \cdots,\quad y_{(n)} = u_1 + \cdots + u_n$

確率変数 $u_1, \cdots, u_n$ の p.e. は単体 $S_n : \{(u_1, \cdots, u_n) : u_\xi \geqq 0,\ \xi = 1, \cdots, n,\ \sum_{\xi=1}^{n} u_\xi \leqq 1\}$ において

(8.7.8)　$$n!\, du_1 \cdots du_n.$$

S_n 以外では，$0 \cdot du_1 \cdots du_n$．これは (7.7.1) で定義された n 変数ディリクレ分布 $D(1, \cdots, 1\,;1)$ の p.e. である．しかし，$u_1, \cdots, u_n$ は 8.7 節の冒頭で定義されたカバーである．しかも，カバーの分布は連続性を仮定した母集団 c.d.f. $F(x)$ に依存しない．このように，カバーは分布に無関係である．

要約すると次を得る．

8.7.4　$(x_{(1)}, \cdots, x_{(n)})$ を連続な c.d.f. $F(x)$ からの標本順序統計量とする．このとき，カバー $u_1 = F(x_{(1)}), u_2 = F(x_{(2)}) - F(x_{(1)}), \cdots, u_n = F(x_{(n)}) - F(x_{(n-1)})$ は n 変数ディリクレ分布 $D(1, \cdots, 1\,; 1)$ を持つ確率変数である．

8.7.4 のカバーの分布は変数に関して完全に対称であることに注意せよ．したがって，カバー $u_1, \cdots, u_n, u_{n+1}$ (ただし $u_{n+1} = 1 - u_1 - \cdots - u_n$) に対応するブロック $B_1^{(1)}, \cdots, B_1^{(n+1)}$ は統計的に同値なブロックということができる．このことと **7.7.2** を用いると，

8.7.5　**8.7.4** における任意の $k(\leqq n)$ 個のカバーは，k 変数ディリクレ分布 $D(1, \cdots, 1\,; n-k+1)$ を持つ．

7.7.4 を用いると

8.7.6　**8.7.4** における任意の $k(\leqq n)$ 個のカバーの和はベータ分布 $Be(k, n-k+1)$ を持つ．

$F(x_{(k)})$ は最初の k 個のカバー $u_1, \cdots, u_k$ の和でもあるから，**8.7.2** は **8.7.6** の特別な

場合であることがわかる．

最後に，**7.7.5** を適用すると，次が成り立つ．

8.7.7　$v_1, \cdots, v_s$ は **8.7.4** におけるカバーの，それぞれ $k_1, \cdots, k_s$ 個の和であるとする．ただし，どのカバーも1つより多くの v に含まれないものとする．このとき，$(v_1, \cdots, v_s)$ の分布は s 変数ディリクレ分布 $D(k_1, \cdots, k_s\,; n-k_1-\cdots-k_s+1)$ になる．

(c) 多次元カバーの標本分布

これまで1次元分布からの標本に関する順序統計量だけを考えてきた．**8.7.4, 8.7.5, 8.7.6, 8.7.7** で述べた結果は，標本ブロックとカバーの定義を2次元，あるいは多次元に拡張しても成り立つ．ここでは2次元の場合を考察しよう．$(x_{1\xi}, x_{2\xi}\,; \xi = 1, \cdots, n)$ を連続な2次元 c.d.f. $F(x_1, x_2)$ からの標本とする．$w = h(x_1, x_2)$ が確率変数となり，しかも連続な c.d.f. $H(w)$ を持つ順序関数 $h(x_1, x_2)$ を導入したとき，$w_\xi = h(x_{1\xi}, x_{2\xi})$, $\xi = 1, \cdots, n$ は分布が $h(x_1, x_2)$ である母集団からの標本を構成し，この標本は順序づけられることになる．標本 $(w_1, \cdots, w_n)$ に対する順序統計量を $(w_{(1)}, \cdots, w_{(n)})$ とする．カバー $u_1', u_2', \cdots, u_{n+1}'$ (ただし $u_{n+1}' = 1 - u_1' - \cdots - u_n'$, $u_1' = H(w_{(1)}), u_2' = H(w_{(2)}) -$

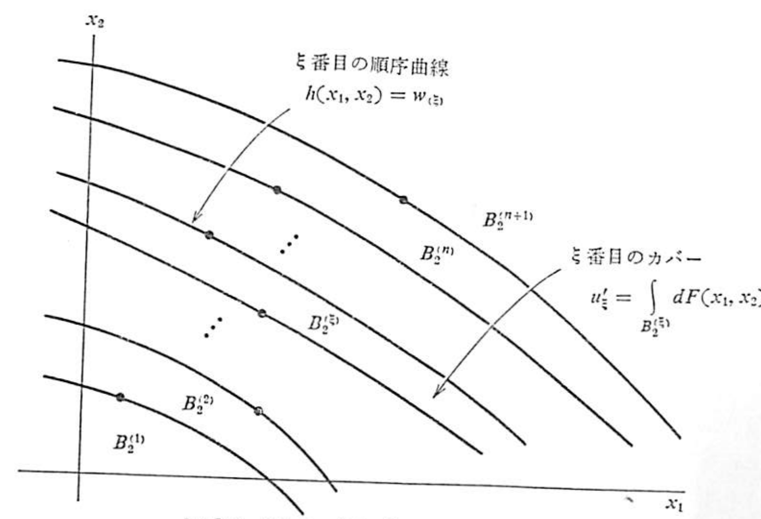

図 8.1　2次元の標本ブロックとカバー．

$H(w_{(1)}), \cdots, u'_n = H(w_{(n)}) - H(w_{(n-1)})$) はそれぞれ 2 次元標本ブロック $B_2^{(1)}, \cdots, B_2^{(n-1)}$ に対応する確率変数である．標本ブロック $B_2^{(1)}, \cdots, B_2^{(n+1)}$ は図 8.1 のように，x_1x_2 平面を曲線 $w_{(\xi)} = h(x_1, x_2)$, $\xi = 1, \cdots, n$ で分割したものである．

このようにすると，k 次元標本ブロック $B_k^{(1)}, \cdots, B_k^{(n+1)}$ に対するカバー $u'_1, \cdots, u'_{n+1}$ を定義する方法も明らかである．もちろん，$B_k^{(1)}, \cdots, B_k^{(n+1)}$ は連続な c. d. f. $H(w)$ を持つ k 変数順序関数 $h(x_1, \cdots, x_k)$ でつくられる．$w_\xi = h(x_{1\xi}, \cdots, x_{k\xi})$, $\xi = 1, \cdots, n$ とする．$(w_1, \cdots, w_n)$ に対応する順序統計量を $(w_{(1)}, \cdots, w_{(n)})$ とすれば，$(w_{(1)}, \cdots, w_{(n)})$ は $h(x_1, \cdots, x_k)$ を分布に持つ母集団からの順序統計量であるから，**8.7.8** が成り立つ．

8.7.8　　順序関数 $h(x_1, \cdots, x_n)$ で定義された k 次元標本ブロックに対するカバー $u'_1, \cdots, u'_n$ の分布は，**8.7.4**, **8.7.5**, **8.7.6**, **8.7.7** で述べた $u_1, \cdots, u_n$ と同じ性質を持つ．

2 次元および多次元のカバーの概念は初期の頃よりもより一般的になっている．矩形標本ブロックを基礎にしてカバーを構成する方法は Wald (1943) が与え，後に Tukey (1947) が一般化している．

ここでは 2 次元の場合について，トゥキーの方法を考えてみよう．n 個の標本点

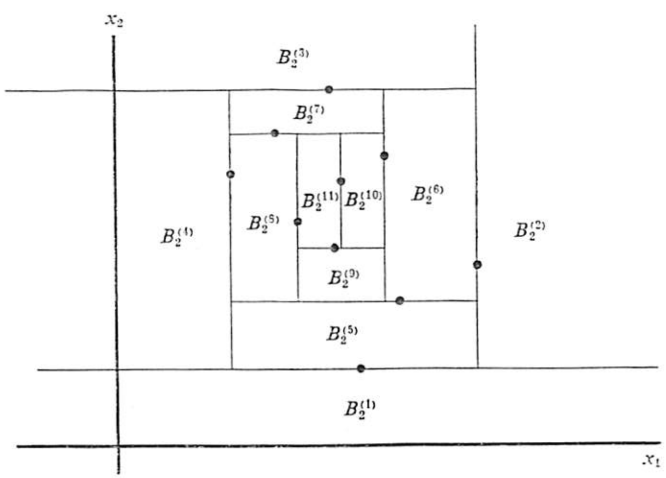

図 8.2　順序関数のグラフとして，水平線と垂線を交互に用いてつくられた 2 次元標本ブロックの例．

$(x_{1\xi}, x_{2\xi})$, $\xi = 1, \cdots, n$ について，ただ 1 つの順序関数 $h(x_1, x_2)$ を用いずに，標本点と同じ個数つまり n 個の順序関数 $h_\eta(x_1, x_2)$, $\eta = 1, \cdots, n$ を用いることができる．

各関数 $h_\eta(x_1, x_2)$ は連続な c. d. f. を持ち，母集団分布の c. d. f. は連続な c. d. f. $F(x_1, x_2)$ であると仮定する．もちろん，$h_\eta(x_1, x_2)$ のいくつか，あるいは，すべてが同じ関数であってもよい．$w_\xi^\eta = h_\eta(x_{1\xi}, x_{2\xi})$ とすれば，η の各値に対して，一組の順序統計量 w_ξ^η, $\xi = 1, \cdots, n$ が定まる．さて x_1x_2 平面を次のように切る：曲線 $w_{(1)}^1 = h_1(x_1, x_2)$ で x_1x_2 平面を 2 つの点集合 $B_2^{(1)}$ と $\overline{B}_2^{(1)}$ に分ける．ただし $B_2^{(1)}$ は $h_1(x_1, x_2) < w_{(1)}^1$ なる (x_1, x_2) の集合である．第 1 のカバー u_1^* を標本ブロック $B_2^{(1)}$ の P 測度 ($F(x_1, x_2)$ で定まる確率) で定義する，次に．$w_{(2)}^2 = h_2(x_1, x_2)$ で $\overline{B}_2^{(1)}$ を $B_2^{(2)}$ と $\overline{B}_2^{(2)}$ に分ける．$B_2^{(2)}$ は $h_2(x_1, x_2) < w_{(2)}^2$ なる $\overline{B}_2^{(1)}$ の点集合である．カバー u_2^* は $B_2^{(2)}$ の P 測度とする．この方法を繰り返すと，$B_2^{(2)} \subset \overline{B}_2^{(1)}, \cdots, B_2^{(n)} \subset \overline{B}_2^{(n-1)}$ となるように，集合 $B_2^{(1)}$, $B_2^{(2)}, \cdots, B_2^{(n)}$ が定義される．残りの集合 $B_2^{(n+1)}$ は $B_2^{(1)} \cup \cdots \cup B_2^{(n)} \cup B_2^{(n+1)} = R_2$ となるように定められる．$B_2^{(1)}, \cdots, B_2^{(n)}$ の P 測度が一般のカバー $u_1^*, \cdots, u_n^*$ である．図 8.2 では，$n = 10$ に対して，順序関数 $h_\eta(x_1, x_2)$, $\eta = 1, \cdots, 10$ として x_1 軸と $(\eta - 1)90°$ の角をなす直線の場合が示されている．

k 変数の場合でも，カバー $u_1^*, \cdots, u_n^*$ は同様に定義される．これを試みよ．カバーに関して次が成り立つ．

8.7.9　　k 次元ブロック (上と同様に定義される) に対するカバー $u_1^*, \cdots, u_n^*$ の分布は，**8.7.4**, **8.7.5**, **8.7.6**, **8.7.7** におけるカバー $u_1, \cdots, u_n$ と同じ性質を持つ．

8.7.9 を証明するには，$u_1^*, \cdots, u_n^*$ が $k = 2$ に対して n 変数ディリクレ分布 $D(1, \cdots, 1; 1)$ を持つことを示せば十分であろう．c. d. f. $F(x_1, x_2)$ から取り出された標本 $(x_{1\xi}, x_{2\xi}; \xi = 1, \cdots, n)$ を O_n で表わす．O_n の p. e. は

$$\prod_{\xi=1}^{n} dF(x_{1\xi}, x_{2\xi}). \tag{8.7.9}$$

$(x_{1*}^{(1)}, x_{2*}^{(1)})$ を n 個の標本要素のうち $h_1(x_1, x_2)$ の値が最小な標本要素とし，$(x_{1\xi_1}^{(1)}, x_{2\xi_1}^{(1)}; \xi_1 = 1, \cdots, n-1)$ を $O_{n-1}^{(1)}$ で表わす．$O_{n-1}^{(1)}$ は O_n から $(x_{1*}^{(1)}, x_{2*}^{(1)})$ をのぞいた $(n-1)$ 個の標本要素からなっている．$(x_{1*}^{(1)}, x_{2*}^{(1)})$ と $(x_{1\xi_1}^{(1)}, x_{2\xi_1}^{(1)}; \xi_1 = 1, \cdots, n-1)$ の p. e. は

$$n \cdot dF(x_{1*}^{(1)}, x_{1*}^{(2)}) \prod_{\xi_1=1}^{n-1} dF(x_{1\xi_1}^{(1)}, x_{2\xi_1}^{(1)}). \tag{8.7.10}$$

さて u_1'' を x_1x_2 平面上の，$h_1(x_1,x_2)<h_1(x_{1*}^{(1)},x_{2*}^{(1)})$ なる集合 $B_2^{(1)}$ のカバーとする．**8.7.2** より，u_1'' の分布はベータ分布 $Be(1,n)$ である．すなわち，u_1'' の p.e. は

(8.7.11)
$$n(1-u_1'')^{n-1}\,du_1''.$$

ただし

(8.7.12)
$$u_1''=\int_{B_2^{(1)}}dF(x_1,x_2),\qquad du_1''=dF(x_{1*}^{(1)},x_{2*}^{(1)}).$$

いま，$(2n-2)$ 次元条件つき確率変数 $(x_{1\xi_1}^{(1)},x_{2\xi_1}^{(2)}\,|\,x_{1*}^{(1)},x_{2*}^{(2)};\xi_1=1,\cdots,n)$ の分布の p.e. は (8.7.10) と (8.7.11) との比であるから

(8.7.13)
$$\prod_{\xi_1=1}^{n-1}dF^{(1)}(x_{1\xi_1}^{(1)},x_{2\xi_1}^{(1)}).$$

ここに

(8.7.14)
$$F^{(1)}(x_{1\xi_1}^{(1)},x_{2\xi_1}^{(1)})=\frac{F(x_{1\xi_1}^{(1)},x_{2\xi_1}^{(1)})}{1-u_1''}.$$

よって，$(x_{1*}^{(1)},x_{2*}^{(1)})$ の値，つまり u_1'' の値が与えられると，もとの標本のうち残りの $(n-1)$ 個の要素は c.d.f.

(8.7.15)
$$F^{(1)}(x_1,x_2)=\frac{F(x_1,x_2)}{1-u_1''}$$

を持つ母集団からの $(n-1)$ 個の標本 $O_{n-1}^{(1)}$ と同じものになる．c.d.f. $F^{(1)}(x_1,x_2)$ は $B_2^{(1)}$ に確率 0 を与え，もとの母集団分布による $\bar{B}_2^{(1)}$ に含まれる割合を 1 に正規化したものである．

この方法を繰り返し，$(x_{1*}^{(2)},x_{2*}^{(2)})$ を $O_{n-1}^{(1)}$ のうち $h_2(x_1,x_2)$ が最小になる $(n-1)$ 個の中の標本要素とし，$(x_{1\xi_2}^{(2)},x_{2\xi_2}^{(2)};\xi_2=1,\cdots,n-2)$ を $O_{n-2}^{(2)}$ とする．$O_{n-2}^{(2)}$ は $O_{n-1}^{(1)}$ から $(x_{1*}^{(2)},x_{2*}^{(2)})$ をのぞいた $(n-2)$ 個の要素である．$(x_{1*}^{(2)},x_{2*}^{(2)})$ と $(x_{1\xi_2}^{(2)},x_{2\xi_2}^{(2)};\xi_2=1,\cdots,n-2)$ の p.e. は

(8.7.16)
$$(n-1)\,dF^{(1)}(x_{1*}^{(2)},x_{2*}^{(2)})\prod_{\xi_2=1}^{n-2}dF^{(1)}(x_{1\xi_2}^{(2)},x_{2\xi_2}^{(2)}).$$

$B_2^{(2)}$ を $\overline{B}_2^{(1)}$ の部分集合で $h_2(x_1,x_2)<h_2(x_{1*}^{(2)},x_{2*}^{(2)})$ なる集合とする．$u_2''=\int_{B_2^{(2)}}dF^{(1)}(x_1,x_2)$ は $F^{(1)}(x_1,x_2)$ によって定まる $B_2^{(2)}$ の条件つきカバーである．**8.7.2** の特殊な場合として，u_2'' の分布は p.e.

(8.7.17)
$$(n-1)(1-u_2'')^{n-2}\,du_2''.$$

$(x_{1\xi_2}^{(2)},x_{2\xi_2}^{(2)}\,|\,x_{1*}^{(2)},x_{2*}^{(2)};\xi_2=1,\cdots,n-2)$ の p.e. は

(8.7.18)
$$\prod_{\xi_2=1}^{n-2}dF^{(2)}(x_{1\xi_2}^{(2)},x_{2\xi_2}^{(2)}).$$

ここに

(8.7.18 *a*)
$$F^{(2)}(x_{1\xi_2}^{(2)},x_{2\xi_2}^{(2)})=\frac{F^{(1)}(x_{1\xi_2}^{(2)},x_{2\xi_2}^{(2)})}{1-u_2''}.$$

よって，$(x_{1*}^{(1)},x_{2*}^{(1)})$，$(x_{1*}^{(2)},x_{2*}^{(2)})$ の値が与えられたとき，$O_{n-2}^{(2)}$ の要素は c.d.f.

(8.7.19)
$$F^{(2)}(x_1,x_2)=\frac{F^{(1)}(x_1,x_2)}{1-u_2''}$$

を持つ母集団からの標本と同じように解釈できる．u_1''，u_2''，$O_{n-2}^{(2)}$ の要素の分布は (8.7.11)，(8.7.17)，(8.7.18) の積である．

この方法を繰り返すと，条件つきカバーの列 $u_1'',\cdots,u_n''$ が得られ，p.e.

(8.7.20)
$$n!(1-u_1'')^{n-1}\cdots(1-u_n'')^{1-1}\,du_1''\cdots du_n''$$

を持つことがわかる．しかし，**8.7.9** で示したカバー u_ξ^*，$\xi=1,\cdots,n$ で表わせば

(8.7.21)
$$\begin{aligned}u_1''&=u_1^*\\ u_2''&=\frac{u_2^*}{1-u_1^*}\\ &\ \ \vdots\\ u_n''&=\frac{u_n^*}{1-u_1^*-\cdots-u_{n-1}^*}.\end{aligned}$$

この変換を，(8.7.20) に用いると

(8.7.22)
$$n!\,du_1^*\cdots du_n^*$$

となる．これは n 変数ディリクレ分布 $D(1,\cdots,1;1)$ の p.e. である．これで **8.7.9** の議論を終わる．

カバーに関する詳細な結果は Fraser (1951, 1953)，Fraser と Guttman (1956)，Kemperman (1956) によって得られている．特に，順序関数を用いて統計的に同値なブロックの構成が詳しく示されている．Tukey (1948) は標本ブロックおよびカバーの概念を，不連続な c.d.f. の場合に拡張している．

8.8 有限母集団からの標本順序統計量

π_N は異なる x 値，たとえば，$x_{o1}<\cdots<x_{oN}$ を持つ有限母集団とする．$(x_{(1)},\cdots,x_{(n)})$ を π_N からの大きさ n の標本順序統計量とする． k 番目の順序統計量 $x_{(k)}$ について考えよう．組合せ理論より

$$p(x_{(k)}=x_{ot})=\frac{\binom{t-1}{k-1}\binom{N-t}{n-k}}{\binom{N}{n}}=p_{N,n,k}(t) \tag{8.8.1}$$

は明らかである． ただし，$t=k,k+1,\cdots,N-n+k$. ここで，$p_{N,n,k}(t)$ には2つの見方がある：（i）確率変数 $x_{(k)}$ の p.f. として，x_{ot} をこの確率変数の質点と考える．$t=k,k+1,\cdots,N-n+k$.　（ii）確率変数 t の p.f. と解釈する． つまり，標本の k 番目の順序統計量が π_N における x 値の階数に等しくなるとき，x 値の階数を表わす確率変数である． 確率変数 t は一連の整数 $k,k+1,\cdots,N-n+k$ に質点を持ち，一方，確率変数 $x_{(k)}$ の質点は $x_{o,k},x_{o,k+1},\cdots,x_{o,N-n+k}$ であるから，$x_{(k)}$ よりも t を取り扱う方が簡単であろう．

t の階乗モーメントは容易に求められる．事実，(3.3.15) の記号を用いれば

$$\mathscr{E}((t+r-1)^{[r]})=\sum_{t=k}^{N-n+k}(t+r-1)^{[r]}p_{N,n,k}(t) \tag{8.8.2}$$

$$=\frac{(k+r-1)^{[r]}\binom{N+r}{n+r}}{\binom{N}{n}}\cdot\sum_{t=k}^{N-n+k}\frac{\binom{t+r-1}{k+r-1}\binom{N-t}{n-k}}{\binom{N+r}{n+r}}.$$

右辺の和は1だから

$$\mathscr{E}((t+r-1)^{[r]})=\frac{(k+r-1)^{[r]}\binom{N+r}{n+r}}{\binom{N}{n}}. \tag{8.8.3}$$

$r=1,2$ とすると，t の平均と分散は

$$\mathscr{E}(t)=\frac{k(N+1)}{(n+1)}$$
$$\sigma^2(t)=\frac{k(N+1)(N-n)(n-k+1)}{(n+1)^2(n+2)}. \tag{8.8.4}$$

まとめると

8.8.1　π_N は相異なる x 値 $x_{o1}<\cdots<x_{oN}$ を持つ有限母集団とする．$(x_{(1)},\cdots,x_{(n)})$ が π_N からの標本順序統計量ならば，$x_{(k)}$ の p.f. は (8.8.1) で与えられ，質点は x_{ot}, $t=k,k+1,\cdots,N-n+k$. $\mathscr{E}((t+r-1)^{[r]})$ の値は (8.8.3) で与えられる．

確率変数の無限列

$$u_{k,N}=\frac{t}{N},\qquad N=k,k+1,\cdots \tag{8.8.5}$$

の極限 $\lim_{N\to\infty}u_{k,N}$ は無限母集団において，x 値が $x_{(k)}$ よりも大きくなっている要素の割合を示す． $\lim_{N\to\infty}u_{k,N}$ は c.d.f. $F(x)$ を持つ無限母集団からの標本の場合の $F(x_{(k)})$ に対応し，**8.7.2** に応じて，ベータ分布 $Be(k,n-k+1)$ を持つことは直観的に明らかである．事実，これを示すには，次のようにする．

$$\lim_{N\to\infty}\mathscr{E}\left(\frac{t}{N}\right)^r=\lim_{N\to\infty}\mathscr{E}\left(\frac{(t+r-1)^{[r]}}{N^r}\right)=\frac{\Gamma(n+1)\Gamma(k+r)}{\Gamma(k)\Gamma(n+r+1)}. \tag{8.8.6}$$

これはベータ分布 $Be(k,n-k+1)$ に従う確率変数の r 次のモーメントである．**5.5.3** より，確率変数 (8.8.5) の分布はベータ分布に近づくことがわかる．したがって

8.8.2　π_N, $N=n,n+1,\cdots$ は異なる x 値 $x_{o1}<\cdots<x_{oN}$ を要素に持つ有限母集団の列とする．$(x_{(1)},\cdots,x_{(n)})$ は π_N からの大きさ n の標本順序統計量とし，$u_{k,N}$ は π_N の中でその x 値が $x_{(k)}$ よりも大きい要素の割合とする． このとき，確率変数列 $u_{k,N}\,;N=n,n+1,\cdots$ の分布はベータ分布 $Be(k,n-k+1)$ に収束する．

問　題

8.1 $(x_1,\cdots,x_n)$ は，平均値のまわりの有限なモーメント $\mu_2,\cdots,\mu_{2r}$ を持つ母集団からの（確率）標本とする．

$$m_r=\frac{1}{n}\sum_{\xi=1}^{n}(x_\xi-\bar{x})^r$$

ならば

$$\sigma^2(m_r)=\frac{1}{n}[\mu_{2r}-\mu_r^2-r(r-1)\mu_2\mu_{r-2}\mu_r+r\mu_{r-1}(r\mu_2\mu_{r-1}-2\mu_{r+1})]+O\left(\frac{1}{n^2}\right)$$

であることを示せ．

8.2 $\bar{x}_1$, $\bar{x}_2$ をそれぞれ $N(\mu_1,\sigma^2)$, $N(\mu_2,\sigma^2)$ からの大きさ n_1, n_2 の独立な標本の標

本平均とし，s_1^2，s_2^2 を標本分散とすれば

$$\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\left[\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\right]^{\frac{1}{2}}}$$

は“スチューデント”の t 分布 $S(n_1+n_2-2)$ に従うことを示せ．

8.3 $(x_{11},\cdots,x_{1n})$，$(x_{21},\cdots,x_{2n})$ はそれぞれ $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$ からの独立な標本とし，$\bar{x}_1$，$\bar{x}_2$ は標本平均，s_1^2，s_2^2 は標本分散，$r=\dfrac{1}{(n-1)s_1s_2}\displaystyle\sum_{\xi=1}^{n}(x_{1\xi}-\bar{x}_1)(x_{2\xi}-\bar{x}_2)$ は標本間の相関係数とする．このとき

$$\frac{[(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)]\sqrt{n}}{[s_1^2+s_2^2-2rs_1s_2]^{\frac{1}{2}}}$$

は“スチューデント” t 分布 $S(n-1)$ を持つことを示せ．

8.4 $\bar{x}_1$，$\bar{x}_2$ は，分散がともに等しく σ^2 になる分布からの，大きさ n_1，n_2 の独立な標本の標本平均とする．$\bar{x}$ をこの標本を合わせた大きさ (n_1+n_2) の標本の標本平均とすれば，$\bar{x}-\bar{x}_1$ の分散は $\sigma^2n_2/[n_1(n_1+n_2)]$ であることを示せ．

8.5 x が分布 $N(\mu,\sigma^2)$ を持つ確率変数で，$F(x)$ がこの正規分布の c.d.f. ならば，x と $F(x)$ の相関係数は $\sqrt{3/\pi}$ であることを示せ．

8.6 (x,y) は相関係数 ρ を持つ任意の2次元正規分布を持つ確率変数とする．y の c.d.f. を $F(y)$ とすれば，x と $F(y)$ との相関係数は $\rho\sqrt{3/\pi}$ になることを示せ．

8.7 $r=2$，$g(x_1,x_2)\equiv x_1x_2$ に対して **8.2.2** を用いて，1次および2次の有限なモーメント μ_1'，μ_2' を持つ分布からの大きさ n の標本について，$Q_{[2]}(x_1,\cdots,x_n)$ をつくれ．$(\mu_1')^2$ の不偏推定量を求めよ．この推定量の分散を計算せよ．

8.8. $(x_1,\cdots,x_n)$ を分散 σ^2 を持つ大きさ N の有限母集団 π_N からの単純確率標本とするとき，任意の実数 $c_1,\cdots,c_n$ に対して，$c_1x_1+\cdots+c_nx_n$ の分散は

$$\left[(c_1^2+\cdots+c_n^2)-\frac{1}{N}(c_1+\cdots+c_n)^2\right]\sigma^2$$

になることを示せ．これを用いて，π_N からの大きさ n_1，n_2 の2つの（確率）標本の標本平均の差の分散は $\sigma^2(1/n_1+1/n_2)$ になることを示せ．ただし $n_1+n_2<N$ で，一度取り出した標本はもとに戻さないとする．

8.9 大きさ $n_1,\cdots,n_k$ の標本は同一の正規分布 $N(\mu,\sigma^2)$ からの独立な標本とする．$\bar{x}_1,\cdots,\bar{x}_k$ をこの標本平均，$s_1^2,\cdots,s_k^2$ を標本分散とする．

$$u=\frac{1}{\sigma^2}[(n_1-1)s_1^2+\cdots+(n_k-1)s_k^2]$$

$$v=\frac{1}{\sigma^2}[n_1(\bar{x}_1-\bar{x})^2+\cdots+n_k(\bar{x}_k-\bar{x})^2]$$

$$w=\frac{1}{\sigma^2}[n(\bar{x}-\mu)^2]$$

と置く．ただし，$\bar{x}$ はこの標本を合わせた大きさ $n=n_1+\cdots+n_k$ の標本の標本平均で

ある．コクランの定理または特性関数を用いて，u，v，w はそれぞれカイ2乗分布 $C(n_1+\cdots+n_k-k)$，$C(k-1)$，$C(1)$ に従う独立な確率変数であることを示せ．

8.10 $(y_1,\cdots,y_n)$ はそれぞれ $N(\mu+\beta x_1,\sigma^2),\cdots,N(\mu+\beta x_n,\sigma^2)$ からの独立な確率変数とする．ただし μ，β，$x_1,\cdots,x_n$ は定数である．$\hat{\mu}$，$\hat{\beta}$ を，平方和

$$\sum_{\xi=1}^{n}(y_\xi-\mu-\beta x_\xi)^2$$

を最小とする μ，β の値とし，かつ

$$v=\frac{1}{\sigma^2}\sum_{\xi=1}^{n}(y_\xi-\hat{\mu}-\hat{\beta}x_\xi)^2$$

とする．$(\hat{\mu},\hat{\beta})$ は2次元正規分布 $N(\{\mu,\beta\},\|\sigma_{ij}\|)$ を持つことを示せ．ここに

$$\|\sigma_{ij}\|=\left\|\begin{matrix}\dfrac{(\sum x_\xi^2)\sigma^2}{na_n} & -\dfrac{\bar{x}\sigma^2}{a_n}\\ -\dfrac{\bar{x}\sigma^2}{a_n} & \dfrac{\sigma^2}{a_n}\end{matrix}\right\|,\qquad a_n=\sum_{\xi}(x_\xi-\bar{x})^2.$$

v はカイ2乗分布 $C(n-2)$ に従い，$(\hat{\mu},\hat{\beta})$ と v は独立であることを示せ．

8.11 $(x_1,\cdots,x_n)$ が $N(\mu,\sigma^2)$ からの標本のとき

$$u=\frac{1}{\sigma^2}\sum_{\xi=1}^{n}(x_\xi-\mu+\delta)^2$$

の特性関数は

$$e^{-\frac{1}{2}\beta^2}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\beta^2}{2}\right)^r(1-2it)^{-\frac{1}{2}n}\exp\left(\frac{\beta^2it}{1-2it}\right).$$

したがって，u の p.d.f. は

$$f(u)=e^{-\frac{1}{2}\beta^2}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\beta^2}{2}\right)^r f_{n+2r}(u)$$

になることを示せ．ただし $\beta^2=n\delta^2/\sigma^2$．$f_{n+2r}(u)$ はカイ2乗分布 $C(n+2r)$ の p.d.f. である．（u の分布は Fisher (1928 b) によってはじめて得られた．これを自由度 n，パラメータ β^2 の**非心カイ2乗分布**という．）

8.12 $N(\mu,\sigma^2)$ からの標本の大きさ，平均，分散をそれぞれ n，$\bar{x}$，s^2 とする．

$$t=\frac{(\bar{x}-\mu+\delta)\sqrt{n}}{s}$$

の分布は p.d.f.

$$f(t)=\frac{e^{-\frac{1}{2}\gamma^2}\Gamma\left(\frac{1}{2}n\right)}{\sqrt{\pi(n-1)}\,\Gamma\left(\frac{1}{2}n-\frac{1}{2}\right)}\left(1+\frac{t^2}{n-1}\right)^{-\frac{1}{2}n}g(t)$$

になることを示せ．ただし

$$g(t)=\sum_{r=0}^{\infty}\left(\frac{\sqrt{2}\gamma t}{\sqrt{n-1}}\right)^r\left(1+\frac{t^2}{n-1}\right)^{-\frac{1}{2}r}\frac{\Gamma\left[\frac{1}{2}(n+r)\right]}{r!\,\Gamma\left(\frac{1}{2}n\right)}$$

かつ $\gamma=\sqrt{n}\,\delta/\sigma$. (Tang (1938) によって最初に得られたこの分布を，自由度 $(n-1)$，パラメータ γ を持つ非心スチューデント分布という.)

8.13　母集団の行要素，列要素，層要素の数がそれぞれ N_1, N_2, N_3 である有限母集団について，(8.6.34)，(8.6.35) に相当する公式を見い出せ.

8.14　$(x_1,\cdots,x_n)$ が矩形分布 $R\left(\frac{1}{2}\theta,\theta\right)$ からの標本のとき，$u=\max(x_1,\cdots,x_n)$ の標本分布は $(0,\theta)$ 上で nu^{n-1}/θ^n，その他で 0 となる p.d.f. を持つことを示せ.

8.15　矩形分布 $R\left(\frac{1}{2},1\right)$ からの大きさ $(2n+1)$ の標本の標本メディアンは，$Be(n+1,n+1)$ なるベータ分布を持つことを示せ.

8.16　$(x_1,\cdots,x_n)$ が矩形分布 $R(\mu,\omega)$ からの標本ならば，標本範囲 r の p.d.f. は $(0,\omega)$ 上では

$$\frac{n(n-1)}{\omega}\left(\frac{r}{\omega}\right)^{n-2}\left(1-\frac{r}{\omega}\right).$$

その他では 0 になることを示せ.

8.17　$(x_{(1)},\cdots x_{(n)})$ は連続な c.d.f. $F(x)$ を持つ母集団からの大きさ n の標本順序統計量とする. $(F(x_{(k_1)}),F(x_{(k_2)}))$, $1\leqslant k_1<k_2\leqslant n$ の共分散行列は

$$\left\|\begin{array}{cc}\dfrac{p_1(1-p_1)}{(n+2)} & \dfrac{p_1(1-p_2)}{(n+2)}\\[2mm] \dfrac{p_1(1-p_2)}{(n+2)} & \dfrac{p_2(1-p_2)}{(n+2)}\end{array}\right\|$$

であることを示せ. ただし $p_1=\dfrac{k_1}{n+1}$, $p_2=\dfrac{k_2}{n+1}$.

8.18　$(x_{(1)},\cdots,x_{(n)})$ が連続な c.d.f $F(x)$ を持つ母集団からの大きさ n の標本順序統計量のとき，標本範囲 $R=x_{(n)}-x_{(1)}$ の平均値は

$$\mathscr{E}(R)=\int_{-\infty}^{\infty}\{1-F^n(x)-[1-F(x)]^n\}\,dx$$

で与えられることを示せ. これは Tippett (1925) による.

8.19（続き）　R の c.d.f. は

$$n\int_{-\infty}^{\infty}\{F(x+R)-F(x)\}^{n-1}\,dF(x)$$

で与えられることを示せ.

8.20　矩形分布 $R(\mu,\omega)$ からの大きさ n の標本において，$x_{(1)}$，$x_{(n)}$ をそれぞれ標本最小要素，標本最大要素とすれば，$r=\left[\frac{1}{2}(x_{(n)}+x_{(1)})-\mu\right]\Big/(x_{(n)}-x_{(1)})$ の p.e. は，区間 $(-\infty,\infty)$ 上で

$$g(r)\,dr=(n-1)(1+2|r|)^{-n}\,dr$$

になることを示せ. [Carlton (1946)].

8.21　$x_{(1)}$，$x_{(n)}$ は連続な c.d.f. $F(x)$ からの大きさ n の標本の標本最小要素，標本最大要素の順序統計量とする. 確率変数 $2n\sqrt{F(x_{(1)})(1-F(x_{(n)}))}$ は $n\longrightarrow\infty$ のとき，平均 $\frac{1}{2}\pi$，分散 $4-\pi^2/4$ の極限分布に持つことを示せ. [Elfving (1947)].

8.22　$(x_{(1)},\cdots,x_{(n)})$ は p.d.f.

$$f(x)=\begin{cases}\dfrac{1}{\sigma}e^{-(x-\mu)/\sigma}, & x>\mu\\ 0, & x\leqslant\mu\end{cases}$$

からの標本順序統計量とする. 確率変数 $t=(x_{(1)}-\mu)/w$ の p.d.f. $g(t)$ は

$$g(t)=\begin{cases}n[1+nt/(n-1)]^{-n}, & t>0\\ 0, & t\leqslant 0\end{cases}$$

で与えられることを示せ. ただし w は

$$w=\frac{1}{n-1}\sum_{\xi=2}^{n}(x_{(\xi)}-x_{(1)})$$

である. [Guttman (1960)].

8.23　$(x_1,\cdots,x_n)$ が $x_1,\cdots,x_n$ に関して対称な p.d.f. $f(x_1,\cdots,x_n)$ を持つ n 次元確率変数であるとき，$(x_1,\cdots,x_n)$ の順序統計量 $(x_{(1)},\cdots,x_{(n)})$ は $-\infty<x_{(1)}<\cdots<x_{(n)}<\infty$ 上では $n!f(x_{(1)},\cdots,x_{(n)})$，その他では 0，になる p.d.f. を持つことを示せ.

8.24　$(x_1,\cdots,x_n)$ は矩形分布 $R\left(\frac{1}{2},1\right)$ からの標本で，

$$y=(x_1\cdots x_n)^{1/n}$$

とすれば，y の p.d.f. は

$$f(y)=\begin{cases}\dfrac{n^n y^{n-1}}{(n-1)!}(-\log y)^{n-1}, & y\in(0,1)\\ 0, & \text{その他}\end{cases}$$

になることを示せ.

8.25　m 組の独立で，大きさ n の標本を矩形分布 $R\left(\frac{1}{2},1\right)$ から抽出する. $z_1,\cdots,z_m$ をそれぞれ各組の標本最大順序統計量とすれば，$u=\prod_{i=1}^{m}z_i$ の p.d.f. は

$$f(u)=\begin{cases}\dfrac{n^m}{(m-1)!}u^{n-1}(-\log u)^{m-1}, & u\in(0,1)\\ 0, & \text{その他}\end{cases}$$

であることを示せ. これは Rider (1955) の結果である.

8.26　$(x_1,\cdots,x_n)$ はガンマ分布 $G(\mu)$ からの標本で，$h(x_1,\cdots,x_n)$ は

$$h(x_1,\cdots,x_n)=h(cx_1,\cdots,cx_n)$$

がすべての $c>0$ に対して成り立つ確率変数とする. このとき $(x_1+\cdots+x_n)$ と $h(x_1,\cdots,x_n)$ が独立になることを示せ. [Pitman (1937 a)].

8.27　$(x_1,\cdots,x_n)$ は p.d.f. $f(x)$ からの標本とし，すべてが 0 ではない任意の定数の組 $c_1,\cdots,c_n$ に対して，比 $(c_1x_1+\cdots+c_nx_n)/(x_1+\cdots+x_n)$ と和 $(x_1+\cdots+x_n)$ が独立ならば，$f(x)$ はガンマ分布の p.d.f. であることを示せ. [Laha (1954)].

8.28　$(x_{(1)},\cdots,x_{(n)})$ は，連続な c.d.f.（有限な平均 μ，分散 σ^2）を持つ母集団からの大きさ n の標本順序統計量であるとする. シュワルツの不等式を用いて

$$\mathscr{E}(x_{(n)}) \leqslant \mu + \frac{(n-1)\sigma}{\sqrt{2n-1}}$$

を示せ．これと類似した型が Hartley と David (1954) によって得られている．

8.29 $(x_{(1)}, \cdots, x_{(n)})$ は連続な c.d.f. $F(x)$ からの大きさ n の標本順序統計量である．このとき確率変数

$$y_i = \left[\frac{F(x_{(i)})}{F(x_{(i+1)})}\right]^i, \qquad i = 1, \cdots, n-1$$

は矩形分布 $R\left(\frac{1}{2}, 1\right)$ に従って，独立に分布することを示せ．[Malmquist (1951), Renyi (1953)].

8.30 $(x_1, \cdots, x_n)$ は矩形分布 $R\left(\frac{1}{2}\omega, \omega+\delta\right)$，$\left(\text{ただし } \omega > 0,\ 0 < \frac{1}{2}\delta < \omega\right)$ からの標本とする．I_ξ は確率区間 $\left(x_\xi - \frac{1}{2}\delta, x_\xi + \frac{1}{2}\delta\right)$ で，E を事象 $\bigcup_{\xi=1}^{n} I_\xi$，$I$ を区間 $(0, \omega)$ とする．

$$\mathscr{E}(E \cap I) = \omega\left[1 - \left(1 - \frac{1}{\omega+\delta}\right)^n\right]$$

を示せ．[Robbins (1944 *a*)].

8.31 長さ L の線分上から"ランダム"に n 個の点をとる．$0 < d < L/(n-1)$ ならば，どの2点も d より近くならない確率は

$$(L - (n-1)d)^n / L^n$$

であることを示せ．[Parzen (1960)].

8.32 $(x_1, \cdots, x_n)$ は平均 μ，分散 σ^2 を持つ正規母集団からの標本，d は

$$d = \frac{1}{n}\sum_{\xi=1}^{n} |x_\xi - \mu|$$

で定義される標本平均偏差とする．

$$\mathscr{E}(d) = \sqrt{2/\pi}\,\sigma, \qquad \sigma^2(d) = \frac{\sigma^2}{n}\left(1 - \frac{2}{\pi}\right)$$

を示せ．

8.33 x は，平均 μ，分散 σ^2 を持ち，p.d.f. $f(x)$ を持つ確率変数である．$\varphi(t)$ を x の特性関数，$\bar{x}$，s^2 を $f(x)$ からの大きさ n の標本の標本平均，標本分散とする．すべての n に対して，$\bar{x}$ と s^2 が独立ならば，$\varphi(t)$ は微分方程式

$$\varphi\frac{d^2\varphi}{dt^2} - \left(\frac{d\varphi}{dt}\right)^2 + \sigma^2\varphi^2 = 0$$

を満たすことを示せ．初期条件 $\varphi(0) = 1$，$\varphi'(0) = i\mu$ の解は $\varphi(t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}$．したがって，$f(x)$ は正規分布 $N(\mu, \sigma^2)$ の p.d.f. になることを示せ．[Lukacs (1942)].

8.34 $x_{(k)}$ は連続な c.d.f. $F(x)$ からの大きさ m の k 番目に小さい標本順序統計量とする．y は同じ c.d.f. からの大きさ n の独立な別の標本 $(x_1, \cdots, x_n)$ のうち，$x_{(k)}$ をこえない標本成分の個数を表わす確率変数であるとする．y の p.f. は

$$\frac{\binom{m}{k}\binom{n}{y}}{\binom{m+n}{k+y}} \cdot \frac{k}{(k+y)}, \qquad y = 0, 1, \cdots, n$$

で，r 次階乗モーメントは

$$\mathscr{E}(y^{[r]}) = \frac{m!\,n!\,(k+r-1)!}{(k-1)!\,(m+r)!\,(n-r)!}, \qquad r \leqslant n$$

で与えられることを示せ．[Epstein (1954)].

8.35（続き）別の標本は $x_{(k)}$ より小なる要素が y 個になるまで抽出されるとする．この標本の大きさ n は確率変数になり，次の p.f.

$$\frac{\binom{n-1}{y-1}\binom{m-1}{k-1}}{\binom{m+n-1}{y+k-1}} \cdot \frac{m}{(m+n)}, \qquad n = y, y+1, \cdots$$

を持つ．

$$\mathscr{E}[(n-y)^{[r]}] = \frac{(y+r-1)!\,(k-r-1)!\,(m-k+r)!}{(y-1)!\,(k-1)!\,(m-k)!}$$

$(r \leqslant k-1)$ を示せ．[Wilks (1959 *b*)].

8.36 $1, 2, \cdots, N$ と印をつけられた N 枚の札からなる有限母集団 π_N を考える．π_N から重複を許さずに抽出された大きさ n の標本において，x を標本最大数とする．x の p.f. は

$$\binom{x-1}{n-1} \Big/ \binom{N}{n}, \qquad x = n, \cdots, N$$

で

$$\mathscr{E}(x) = \frac{n}{n+1}(N+1), \qquad \sigma^2(x) = \frac{n(N-n)(N+1)}{(n+1)^2(n+2)}$$

を示せ．

8.37（続き）大きさ $(2n+1)$ の標本を有限母集団 π_N からもどさずに抽出する．y を標本メディアンとすれば，y の p.f. は

$$\binom{y-1}{n}\binom{N-y}{n} \Big/ \binom{N}{2n+1}, \qquad y = n+1, \cdots, N-n$$

で

$$\mathscr{E}(y) = \frac{N+1}{2}, \qquad \sigma^2(y) = \frac{(N-2n-1)(N+1)}{8n+12}$$

になることを示せ．

8.38（続き）s，t を π_N からの大きさ n の標本のうち，それぞれ k_1，k_2 $(k_1 < k_2)$ 番目に小さい標本順序統計量とする．(s, t) の p.f. が

$$\binom{s-1}{k_1-1}\binom{t-s-1}{k_2-k_1-1}\binom{N-t}{n-k_2} \Big/ \binom{N}{n}$$

になり

$$\mathscr{E}(s) = k_1\left(\frac{N+1}{n+1}\right)$$
$$\sigma^2(s) = \frac{k_1(N-n)(N+1)(n-k_1+1)}{(n+1)^2(n+2)}$$
$$\operatorname{cov}(s,t) = \frac{k_1(n-k_2+1)(N-n)(n+1)}{(n+1)^2(n+2)}$$

を示せ.

8.39 $p(x)$, $q(x)$ は同じ質点 $x_1, \cdots, x_k$ を持つ2つの離散型分布である. O_m, O_n をそれぞれ $p(x)$, $q(x)$ からの大きさ m, n の独立な標本とする. $(m_1, \cdots, m_k)$ は O_m の成分のうち, それぞれ $x_1, \cdots, x_k$ なる値を持つ成分の数とする. もちろん, $m_1 + \cdots + m_k = m$ である. 同様に, $(n_1, \cdots, n_k)$ を定義する. $r_{m,n}$ を $(\sqrt{m_1/m}, \cdots, \sqrt{m_k/m})$ と $(\sqrt{n_1/n}, \cdots, \sqrt{n_k/n})$ の相関係数とする. $p(x) \equiv q(x)$ ならば, 任意の $\lambda > 0$ に対して

$$P\left(r_{m,n} > 1 - \frac{\lambda}{2}\right) \geqslant 1 - \frac{k-1}{\lambda}\left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}}\right)^2$$

を示せ. これは Matusita (1957) による.

8.40 (続き) $\sum_{i=1}^{k}(\sqrt{p(x_i)} - \sqrt{q(x_i)})^2 \geqslant \delta^2$ (ただし $\delta > \lambda > 0$) ならば

$$P\left(r_{m,n} < 1 - \frac{\lambda}{2}\right) \geqslant 1 - \frac{k-1}{(\delta-\lambda)^2}\left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}}\right)^2$$

であることを示せ. これも Matusita (1957) の結果である.

8.41 8.3(a) 節における z の c.d.f. は

$$\begin{aligned} F_n(z) &= \frac{1}{n!}\left[z^n - \binom{n}{1}(z-1)^n + \binom{n}{2}(z-2)^n - \cdots + (-1)^k\binom{n}{k}(z-k)^n\right], \\ &\qquad k < z \leqslant k+1, \quad k = 0, 1, \cdots, n-1 \\ &= 0, \quad z \leqslant 0 \\ &= 1, \quad z > n \end{aligned}$$

になることを示せ.

8.42 $(x_{(1)}, \cdots, x_{(n)})$ は矩形分布 $R\left(\frac{1}{2}, 1\right)$ からの大きさ n の標本順序統計量で, $(u_1, \cdots, u_n)$ を区分 $(x_{(1)}, x_{(2)} - x_{(1)}, \cdots, x_{(n)} - x_{(n-1)})$ とすれば, $v = \max(u_1, \cdots, u_n)$ の c.d.f. は

$$\begin{aligned} F_n(v) &= \left[1 - \binom{n}{1}(1-v)^n + \binom{n}{2}(1-2v)^n - \cdots + (-1)^k\binom{n}{k}(1-kv)^n\right], \\ &\qquad \frac{1}{k+1} < v \leqslant \frac{1}{k}, \quad k = n-1, n-2, \cdots, 1 \\ &= 0, \quad v \leqslant 0 \\ &= 1, \quad v > 1 \end{aligned}$$

であることを示せ. (前問を用いて解ける).

8.43 (続き) **区間 $(0,1)$ 上の n 個のランダムな点でつくられた最大区分の分布**

区間 $(0,1)$ 上の順序統計量 $(x_{(1)}, \cdots, x_{(n)})$ はこの区間を $(n+1)$ 個の互いに素な区分

$u_1, \cdots, u_{n+1}$ に分割する. ただし $u_1 + \cdots + u_{n+1} = 1$. $w = \max(u_1, \cdots, u_{n+1})$ の c.d.f. は

$$\begin{aligned} F_n(w) &= \left[1 - \binom{n+1}{1}(1-w)^n + \binom{n+1}{2}(1-2w)^n - \cdots \right. \\ &\qquad \left. + (-1)^k\binom{n+1}{k}(1-kw)^n\right], \\ &\qquad \frac{1}{k+1} < w \leqslant \frac{1}{k}, \quad k = n, n-1, \cdots, 1 \\ F_n(w) &= 0, \quad w \leqslant \frac{1}{n+1} \\ &= 1, \quad w > 1 \end{aligned}$$

で与えられることを示せ.

8.44 $(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)})$ は $N(0,1)$ からの大きさ4の標本順序統計量とするとき, $y = \frac{1}{2}(x_{(3)} + x_{(4)}) - \frac{1}{2}(x_{(1)} + x_{(2)})$ の c.d.f. $G(y)$ は

$$G(y) = 8\left[\frac{1}{\sqrt{2\pi}}\int_0^y e^{-\frac{1}{2}x^2}\,dx\right]^3, \qquad y \geqslant 0$$

であることを示せ. [Walsh (1946)].

8.45 $(x_{(1)}, x_{(2)})$ を $N(\mu, \sigma^2)$ からの大きさ2の標本順序統計量とするとき, この確率変数の平均値は $(\mu - \sigma/\sqrt{\pi}, \mu + \sigma/\sqrt{\pi})$ であることを示せ.

第9章　大標本に対する漸近的標本論

第8章では，大きさ n の有限標本に基づいて標本論を展開してきた．本章では，$n \longrightarrow \infty$ のとき，標本分布に関してどのような結果が得られるかを考察しよう．すなわち，ここでは，c.d.f. $F(x)$ を持つ母集団からの確率過程 $(x_1, x_2, \cdots)$ を対象にする．ただし，この確率過程の要素は互いに独立で，n 個の要素からなる部分集合は c.d.f. $F(x)$ を持つ母集団からの大きさ n の標本である．この確率過程は**無限母集団からの単純無作為標本**，あるいは，**確率分布からの無作為標本**と呼ばれている．本章では，大標本の場合の標本関数の漸近的分布を求めるのに有用な極限定理や結果を導く．

9.1　標本平均の確率収束

大標本の標本平均に関する最も簡単な結果は，Khintchine (1929) の次の定理にまとめられる．

9.1.1　$(x_1, \cdots, x_n)$ が，$\mathscr{E}(x)$ として有限な値 μ を持つ c.d.f. $F(x)$ からの標本であれば，$n \longrightarrow \infty$ のとき，標本平均 $\bar{x}$ は μ に確率収束する．

この証明は，平均 μ が存在すれば，(5.1.10) より $F(x)$ に対応する特性関数 $\varphi(t)$ が

(9.1.1)
$$\varphi(t) = 1 + i\mu t + o(t)$$

なる型に書かれることに注意すればよい．**8.3.2** を用いると，$\bar{x}$ の特性関数は $[\varphi(t/n)]^n$ だから

(9.1.2)
$$\left[\varphi\left(\frac{t}{n}\right)\right]^n = \left[1 + \frac{\mu i t}{n} + o\left(\frac{t}{n}\right)\right]^n.$$

ただし，$n \longrightarrow \infty$ のとき，$n \cdot o(t/n)$ は任意の t に対して 0 に近づく．$n \longrightarrow \infty$ とすれば

(9.1.3)
$$\lim_{n\to\infty}\left[\varphi\left(\frac{t}{n}\right)\right]^n = e^{\mu i t}.$$

これは (2.3.4) で定義された退化 c.d.f. $\varepsilon(x-\mu)$ の特性関数である．**5.4.1** を適用すれば，$n \longrightarrow \infty$ ならば $\bar{x}$ の分布は $\varepsilon(x-\mu)$ に収束する．これは $\bar{x}$ が $n \longrightarrow \infty$ のとき μ に確率収束することと同値である．よって **9.1.1** が証明された．

x の分散 σ^2 と平均 μ が共に存在して，**9.1.1** における母集団分布に関して有限であるならば，チェビシェフの不等式 (3.3.5) により，$\bar{x}$ が μ に確率収束することが簡単に証明される．なぜなら **8.2.1** から $\mathscr{E}(\bar{x}) = \mu$，$\sigma^2(\bar{x}) = \sigma^2/n$ であり，(3.3.5) を用いると，任意の $\varepsilon > 0$ に対して

(9.1.4)
$$P(|\bar{x}-\mu| \geqslant \varepsilon) = P\left(|\bar{x}-\mu| \geqslant \left(\frac{\varepsilon}{\sigma_{\bar{x}}}\right)\cdot\sigma_{\bar{x}}\right) \leqslant \frac{\sigma_{\bar{x}}^2}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

ゆえに

(9.1.5)
$$\lim_{n\to\infty} P(|\bar{x}-\mu| \geqslant \varepsilon) = 0$$

となり，$\bar{x}$ は μ に確率収束する．

9.1.1 は k 次元の場合に拡張される．この公式と証明は読者自身で考えよ．

定理 **9.1.1** は**大数の弱法則**と呼ばれている．これは $n \longrightarrow \infty$ のとき，$\bar{x}$ の確率分布に関して 100% の点が，$\bar{x} = \mu$ を含む任意の近傍に集積して行くことを示しているといえよう．これを保障する条件が（有限な）平均値 μ の存在である．

コーシー分布　一見，整った型に分布しているように思われるが，実は上に述べた確率収束の性質を持たない標本平均の例として **Cauchy** (1853) 分布がある．この p.d.f. は

(9.1.6)
$$f(x) = \frac{\theta_2}{\pi[\theta_2^2 + (x-\theta_1)^2]}.$$

ただし x の範囲は $(-\infty, +\infty)$ で，θ_1，θ_2 は実数で，$\theta_2 > 0$. (9.1.6) に対応する特性関数は

(9.1.7)
$$\varphi(t) = e^{\theta_1 i t - \theta_2 |t|}.$$

したがって，大きさ n の標本平均 $\bar{x}$ の特性関数は

(9.1.8)
$$\left[\varphi\left(\frac{t}{n}\right)\right]^n = [e^{\theta_1 i t/n - \theta_2 |t/n|}]^n = \varphi(t).$$

つまり $\bar{x}$ は x と同一の特性関数を持つ．よって，任意の n に対して $\bar{x}$ の標本分布は母集団の x の分布に等しい．読者は $f(x)$ が $x = \theta_1$ に関して対称ではあるが，θ_1 は分布の平均ではなく，平均が存在しないことを確かめよ．高次のモーメントも存在しない．しかし，θ_1 は分布のメディアンでもあり，分布の対称軸の中心でもある．一方，θ_2 は4

分位区間，すなわち4分位 $\underline{x}_{0.25}$ と $\underline{x}_{0.75}$ の間の距離を表わしている．

9.2　標本和と標本平均の極限分布

(a)　1次元の場合

母集団分布の平均ばかりでなく，その分散も存在するならば，$n \longrightarrow \infty$ のとき，$\bar{x}$ の分布がどうなるかがわかる．これに関する基本的な定理は Lindeberg (1922) による．

9.2.1　z, $\bar{x}$ は，有限な分散 σ^2，平均 μ を持つ分布からの大きさ n の標本の標本和および標本平均である．このとき

$$\lim_{n\to\infty} P\left[\frac{z-n\mu}{\sqrt{n}\,\sigma} \leqslant y\right] = \lim_{n\to\infty} P\left[\frac{\sqrt{n}\,(\bar{x}-\mu)}{\sigma} \leqslant y\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{1}{2}u^2}\,du. \tag{9.2.1}$$

(9.2.1) が成立するとき，便宜上，z は大きな n に対して $N(n\mu, n\sigma^2)$ に従って，**漸近的に正規分布をしている**[*)]（$\bar{x}$ は大きな n に対して $N(\mu, \sigma^2/n)$ に従って，漸近的に正規分布をしている）という．

$(z-n\mu)/\sqrt{n}\,\sigma$ は確率変数 $\sqrt{n}\,(\bar{x}-\mu)/\sigma$ を書き換えたに過ぎないから，前者を示せば十分である．$(z-n\mu)/\sqrt{n}\,\sigma$ の特性関数を $\varphi_n(t)$ とすると

$$\varphi_n(t) = \mathscr{E}\left[e^{it(z-n\mu)/\sqrt{n}\,\sigma}\right] = \mathscr{E}\left[\exp\left(it\sum_{\xi=1}^{n}(x_\xi-\mu)/\sqrt{n}\,\sigma\right)\right] = [\varphi(t)]^n. \tag{9.2.2}$$

ここで $\varphi(t)$ は $(x-\mu)/\sqrt{n}\,\sigma$ の特性関数である．(5.1.10) より

$$\varphi(t) = 1 + i\mathscr{E}\left(\frac{x-\mu}{\sqrt{n}\,\sigma}\right)t - \frac{1}{2}\mathscr{E}\left(\frac{x-\mu}{\sqrt{n}\,\sigma}\right)^2 t^2 + o\left(\frac{t^2}{n}\right). \tag{9.2.3}$$

ただし，任意の $t \neq 0$ に対して，$n\cdot o(t^2/n) \longrightarrow 0\ (n \longrightarrow \infty)$．$t=0$ のとき，$o(t^2/n)=0$．

したがって，任意の t に対して

$$\lim_{n\to\infty}\varphi_n(t) = \lim_{n\to\infty}\left[1-\frac{t^2}{2n}+o\left(\frac{t^2}{n}\right)\right]^n = e^{-t^2/2}. \tag{9.2.4}$$

しかし，**7.2.1** により，$e^{-t^2/2}$ は正規分布 $N(0,1)$ の特性関数であるから，**5.4.1** を用いれば，$n \longrightarrow \infty$ のとき $(z-n\mu)/\sqrt{n}\,\sigma$ の c.d.f. は分布 $N(0,1)$ の c.d.f. に収束する．これで **9.2.1** が成り立つ．

次にド・モアブル＝ラプラスの定理として知られている **9.2.1** の重要な系を述べておく．

*) 大きな n に対して，漸近的分布 $N(n\mu, n\sigma^2)$ を持つということもある．（訳注）

9.2.1 a　x が2項分布 $Bi(n,p)$ を持つ確率変数のとき，z は漸近的に $N(np, npq)$ に従う分布となる．

8.3.3 a で $m=1$ とすれば，2項分布 $Bi(1,p)$ から取り出された大きさ n の標本和 z は2項分布 $Bi(n,p)$ を持つ．**9.2.1** を適用すれば，**9.2.1 a** が得られる．**9.2.1 a** の結果は De Moivre (1718) によって推測されていたが，Laplace (1814) までの1世紀間，明らかにされなかった．Gauss (1809 *b*) は誤差論との関連において，**9.2.1** の近似式を導いている．

(b)　中心極限定理

定理 **9.2.1** は中心極限定理といわれる一般的な結果の特別な場合である．この定理はある条件の下で

$$\lim_{n\to\infty} P\left[\frac{\sum_{\xi=1}^{n}(x_\xi-\mu_\xi)}{\sqrt{\sum_{\xi=1}^{n}\sigma_\xi^2}} \leqslant y\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{1}{2}u^2}\,du \tag{9.2.5}$$

が成り立つことである．ただし $(x_1, x_2, \cdots)$ は平均 $(\mu_1, \mu_2, \cdots)$，分散 $(\sigma_1^2, \sigma_2^2, \cdots)$ を持つ独立確率変数列である．この主張が成立するための条件が Chebyshev (1890)，Feller (1935)，Lévy (1935)，Lindeberg (1922)，Lyapunov (1900, 1901)，Markov (1900) 等により研究されてきた．中心極限定理およびこれに関する問題は Gnedenko と Kolmogorov (1954) により包括的に解明されている．一般中心極限定理の近代的な型は次のように述べられている．

9.2.2　$(x_1, x_2, \cdots)$ は平均 $(\mu_1, \mu_2, \cdots)$，分散 $(\sigma_1^2, \sigma_2^2, \cdots)$，c.d.f. $(F_1(x), F_2(x), \cdots)$ を持つ独立確率変数列とする．$z_n = x_1+\cdots+x_n$，$\zeta_n = \mu_1+\cdots+\mu_n$，$\tau_n^2 = \sigma_1^2+\cdots+\sigma_n^2$，$\tau_n^{*2} = \sigma_1^{*2}+\cdots+\sigma_n^{*2}$ とする．ただし，任意の $\varepsilon>0$ に対して，

$$\sigma_\xi^{*2} = \int_{\mu_\xi-\varepsilon\tau_n}^{\mu_\xi+\varepsilon\tau_n}(x-\mu_\xi)^2\,dF_\xi(x), \qquad \xi = 1,\cdots,n \tag{9.2.6}$$

である．

$$\lim_{n\to\infty} P\left(\frac{z_n-\zeta_n}{\tau_n} \leqslant y\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{1}{2}u^2}\,du \tag{9.2.7}$$

であるための必要十分条件は任意の $\varepsilon > 0$ に対して

(9.2.8)
$$\lim_{n\to\infty} \frac{\tau_n^{*2}}{\tau_n^2} = 1$$

かつ $\lim_{n\to\infty} \tau_n^2 = \infty$ が成り立つことである．

この証明は相当長くなる．詳しくは確率論の本を参照されたい．ここでは省略しよう．条件 (9.2.8) の十分性は Lindeberg (1922)，必要性は Feller (1935) により証明された．

読者は，$(x_1, x_2, \cdots)$ が分散 σ^2 の同一分布を持つ独立確率変数列のとき，(9.2.8) が満たされることを容易に確かめることができるだろう．

(c)　k 次元の場合

9.2.1 の k 次元への拡張は，次のようになる．

9.2.3　$(x_{1\xi}, \cdots, x_{k\xi}\,;\xi = 1, \cdots, n)$ は有限な平均 μ_i, $i = 1, \cdots, k$, 正定値共分散行列 $\|\sigma_{ij}\|$, $i, j = 1, \cdots, k$ を持つ k 変数分布からの大きさ n の標本とする．$(z_1, \cdots, z_k)$, $(\bar{x}_1, \cdots, \bar{x}_k)$ を 8.1 節で定義した標本和，標本平均とすれば，$(z_1, \cdots, z_k)$, $(\bar{x}_1, \cdots, \bar{x}_k)$ はそれぞれ k 変数分布 $N(\{n\mu_i\}, \|n\sigma_{ij}\|)$, $N\left(\{\mu_i\}, \left\|\frac{\sigma_{ij}}{n}\right\|\right)$ を漸近的分布として持つ．すなわち

(9.2.9)
$$\begin{aligned}\lim_{n\to\infty} P\left(\frac{(z_i - n\mu_i)}{\sqrt{n}} \leqslant y_i, i = 1, \cdots, k\right) &= \lim_{n\to\infty} P((\bar{x}_i - \mu_i)\sqrt{n} \leqslant y_i, i = 1, \cdots, k) \\ &= \frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}} \int_{-\infty}^{y_k} \cdots \int_{-\infty}^{y_1} \exp\left(-\frac{1}{2}\sum_{i,j=1}^{k} \sigma^{ij} u_i u_j\right) du_1 \cdots du_k.\end{aligned}$$

9.2.3 は **9.2.1** の議論を直接 k 次元に拡張すれば，容易に成り立つことがわかる．ここで重要なことは，k 次元確率変数 $[(z_1 - n\mu_1)/\sqrt{n}, \cdots, (z_k - n\mu_k)/\sqrt{n}]$ の特性関数 $\varphi_n(t_1, \cdots, t_k)$ を求めて

(9.2.10)
$$\lim_{n\to\infty} \varphi_n(t_1, \cdots, t_k) = \exp\left(-\frac{1}{2}\sum_{i,j=1}^{k} \sigma_{ij} t_i t_j\right)$$

を示すことである．**5.4.1** を k 次元の場合に適用すれば，(9.2.9) がいえる．(9.2.10) はすぐ成り立つ．読者はこれを確かめよ．

9.2.3 の重要な応用の1つとして，ド・モアブルの定理 **9.2.1a** の k 次元への拡張がある．これは次のようになる．.

9.2.3a　$(x_{1\xi}, \cdots, x_{k\xi}\,;\xi = 1, \cdots, n)$ を多項分布 $M(1\,;p_1, \cdots, p_k)$ からの大きさ n の標本とすれば，標本和 $(z_1, \cdots, z_k)$ は，大きな n に対して，漸近分布として $N(\{np_i\}, \|n(p_i\delta_{ij} - p_ip_j)\|)$ を持つ．ただし，δ_{ij} はクロネッカーの記号である．

9.2.3 の系としての **9.2.3a** の証明は，(6.3.3) で与えられた多項分布 $M(1\,;p_1, \cdots, p_k)$ の p.f. から

(9.2.11)
$$\mathscr{E}(x_i) = p_i, \qquad \sigma^2(x_i, x_j) = \sigma_{ij} = p_i\delta_{ij} - p_ip_j$$

を確かめればすぐわかる．

9.3　標本平均の関数の漸近的分布

(a)　一　般　型

標本平均 $\bar{x}$ のある関数，たとえば $g(\bar{x})$，の大きな n に対する漸近的分布を知る必要がしばしばある．この種の問題に関する有用な結果を述べておこう．

9.3.1　$(x_1, \cdots, x_n)$ は共に有限な平均 μ, 分散 σ^2 を持つ分布からの標本とする．$g(x)$ は $x = \mu$ のある近傍で1次導関数 $g'(x)$ を持ち，$g'(\mu) \neq 0$ とする．このとき $g(\bar{x})$ は大きな n に対する漸近的分布として，$N(g(\mu), [\sigma g'(\mu)]^2/n)$ を持つ．

これを証明しよう．$V(\mu)$ を，$V(\mu)$ のすべての点 x で $g'(x)$ が存在するような $x = \mu$ の近傍とする．標本が取り出された母集団の c.d.f. は有限な平均 μ, 分散 σ^2 を持つので，**4.6.1** から，任意の $\varepsilon > 0$ に対して，n_ε が存在して，

(9.3.1)
$$P(\bar{x} \in V(\mu),\ \text{すべての}\ n > n_\varepsilon) > 1 - \varepsilon.$$

$V(\mu)$ の任意の点 $x = \bar{x}$ については

(9.3.2)
$$g(\bar{x}) = g(\mu) + g'(x^*)\cdot(\bar{x} - \mu)$$

と書ける．ただし x^* は $|x^* - \mu| \leqslant |\bar{x} - \mu|$ となる確率変数である．(9.3.2) は

(9.3.3)
$$(g(\bar{x}) - g(\mu))\sqrt{n} = g'(x^*)(\bar{x} - \mu)\sqrt{n}$$

となる．左辺の確率変数を u_n，右辺を v_n と表わせば，(9.3.3) が成立する確率は，

$n > n_\varepsilon$ なるすべての n に対して $1-\varepsilon$ をこえる．ε は任意の正の数だから，確率変数列 $(u_1, u_2, \cdots)$ と $(v_1, v_2, \cdots)$ は一方が法則収束すれば，共に法則収束する．σ^2 は有限だから，**9.2.1** より，確率変数列 $\sqrt{n}(\bar{x}-\mu)$，$n=1,2,\cdots$ は正規分布 $N(0,\sigma^2)$ に法則収束することがわかる．$g'(x)$ は $V(\mu)$ のすべての点で存在するから，$V(\mu)$ したがって $x=\mu$ で連続である．**4.3.7** から，$g'(x^*)$ は $g'(\mu)$ に確率収束する．したがって

(9.3.4) $$\lim_{n\to\infty} P(v_n \leqslant w) = P(g'(\mu)s \leqslant w).$$

ただし s は分布 $N(o,\sigma^2)$ を持つ．よって

(9.3.5) $$\lim_{n\to\infty} P(u_n \leqslant w) = P(t \leqslant w).$$

ここに t は分布 $N(0,(\sigma g'(\mu))^2)$ を持つ．しかし (9.3.5) は，大きな n に対して $g(\bar{x})$ の漸近的分布が $N(g(\mu),[\sigma g'(\mu)]^2/n)$ であることと同値である．これで証明が終わる．

9.3.1 を k 次元の場合に拡張すると，次のようになる．

9.3.1 a　$(x_{1\xi},\cdots,x_{k\xi},\ \xi=1,\cdots,n)$ は有限な平均 $\{\mu_i\}$，正定値共分散行列 $\|\sigma_{ij}\|$，$i,j=1,\cdots,k$ を持つ k 次元分布からの標本とする．$g(x_1,\cdots,x_k)$ は $(\mu_1,\cdots,\mu_k)$ のある近傍で1次偏導関数 $\partial g/\partial x_i = g_i$，$i=1,\cdots,k$ を持つ関数で，$g_i^0 = g_i(\mu_1,\cdots,\mu_k)$ とする．このとき，少なくとも1つの g_i^0 が 0 でなければ，$g(\bar{x}_1,\cdots,\bar{x}_k)$ は，大きな n に対して漸近的分布 $N\Big(g(\mu_1,\cdots,\mu_k),\ \dfrac{1}{n}\sum_{i,j=1}^{n}\sigma_{ij}g_i^0g_j^0\Big)$ を持つ．

この証明は **9.3.1** と同様になされる．練習問題として残しておこう．

(b)　標本平均および標本和の2次関数

(9.2.9) により，独立な k 次元確率変数列

(9.3.6) $$\left(\frac{z_1-n\mu_1}{\sqrt{n}},\cdots,\frac{z_k-n\mu_k}{\sqrt{n}}\right),\qquad n=1,2,\cdots$$

あるいは，これと同値な独立な k 次元確率変数列

(9.3.7) $$((\bar{x}_1-\mu_1)\sqrt{n},\cdots,(\bar{x}_k-\mu_k)\sqrt{n}),\qquad n=1,2,\cdots$$

が k 次元正規分布 $N(\{0\},\|\sigma_{ij}\|)$ に法則収束することがわかる．便宜上，十分大きな n に対して，確率変数 (9.3.6) [あるいは (9.3.7)] の c.d.f. を $F_n(y_1,\cdots,y_k)$，分布 $N(\{0\},\|\sigma_{ij}\|)$ の c.d.f. を $\Phi(y_1,\cdots,y_k)$ で表わそう．$\Phi(y_1,\cdots,y_k)$ は (9.2.9) で与えられる．

7.8.2 では $(x_1,\cdots,x_k)$ が分布 $N(\{\mu_i\},\|\sigma_{ij}\|)$ を持てば，$\sum_{i,j=1}^{k}\sigma^{ij}(x_i-\mu_i)(x_j-\mu_j)$ はカイ2乗分布 $C(k)$ を持つことが示された．2次形式 Q_n を

(9.3.8) $$Q_n = \sum_{i,j=1}^{k}\sigma^{ij}\left(\frac{z_i-n\mu_i}{\sqrt{n}}\right)\left(\frac{z_j-n\mu_j}{\sqrt{n}}\right) = n\sum_{i,j=1}^{k}\sigma^{ij}(\bar{x}_i-\mu_i)(\bar{x}_j-\mu_j)$$

とすれば，確率変数 Q_n，$n=1,2,\cdots$ は $n\longrightarrow\infty$ のとき，カイ2乗分布 $C(k)$ に法則収束する．なぜなら，確率過程 (9.3.6) は k 次元正規分布 $N(\{0\},\|\sigma_{ij}\|)$ に法則収束するので，**4.3.4**，**4.3.6** を適用すると，$Q_1, Q_2,\cdots$ はカイ2乗分布 $C(k)$ に法則収束する．まとめると次のようになる．

9.3.2　$(x_{1\xi},\cdots,x_{k\xi};\xi=1,\cdots,n)$ を有限な平均 μ_i，$i=1,\cdots,k$，有限な正定値共分散行列 $\|\sigma_{ij}\|$ を持つ k 次元分布からの標本とする．(9.3.8) で定義された Q_n は，大きな n に対して漸近的分布としてカイ2乗分布 $C(k)$ を持つ．すなわち，

(9.3.9) $$\lim_{n\to\infty}P(Q_n\leqslant y) = \frac{1}{2\Gamma\left(\frac{1}{2}k\right)}\int_0^y\left(\frac{u}{2}\right)^{\frac{1}{2}k-1}e^{-\frac{1}{2}u}\,du$$

が成り立つ．

$\bar{x}_1,\cdots,\bar{x}_k$ の関数 $Q_n(\bar{x}_1,\cdots,\bar{x}_k)$ は **9.3.1 a** での，点 $(\mu_1,\cdots,\mu_k)$ で1次導関数のうち少なくとも1つが0でないという条件に反している．しかし，この場合，Q_n は $n\longrightarrow\infty$ のとき，極限分布 $C(k)$ を持つ．

(c)　ピヤソンのカイ2乗適合性基準

9.3.2 の重要な例として，標本が多項分布 $M(1;p_1,\cdots,p_k)$ なる k 次元母集団から取り出された場合が考えられる．この分布の平均と分散は (9.2.11) で与えられる．このとき，

(9.3.10) $$\|\sigma^{ij}\| = \|p_i\delta_{ij}-p_ip_j\|^{-1} = \left\|\frac{\delta_{ij}}{p_i}+\frac{1}{p_{k+1}}\right\|$$

を得る．(9.3.8) の σ^{ij} に $\|\delta_{ij}/p_i+1/p_{k+1}\|$，$\mu_i$ に p_i を代入すると

(9.3.11) $$Q_n = \sum_{i=1}^{k+1}\frac{(z_i-np_i)^2}{np_i},\quad ただし\ z_{k+1} = \sum_{\xi=1}^{n}x_{k+1\xi} = n-\sum_{i=1}^{k}z_i.$$

ゆえに **8.3.3 b**，すなわち多項分布 $M(1;p_1,\cdots,p_k)$ からの大きさ n の標本和 $(z_1,\cdots,z_k)$ は分布 $M(n;p_1,\cdots,p_k)$ に従うことに注意すれば，**9.3.2** の系として，次を得る．

9.3.2 a　$(z_1,\cdots,z_k)$ が多項分布 $M(n;p_1,\cdots,p_k)$ を持つ k 次元確率変数ならば

(9.3.12) $$Q_n=\sum_{i=1}^{k+1}\frac{(z_i-np_i)^2}{np_i}$$

は，大きな n に対して漸近的にカイ2乗分布 $C(k)$ に従う．

(9.3.12) は標本"頻度" z_i とこの平均値との差の2乗を平均値で重みづけした量の和である．これは**カイ2乗適合性基準**といわれ，K. Pearson (1900) が最初に紹介した．この量は z_i がおのおのの平均値からどのくらい離れているかを示す指標と考えられる．特定の（大）標本に対する指標の大きさは，その漸近的分布つまりカイ2乗分布 $C(k)$ から計算される確率で表わされる．この問題を深く考察すると，統計的仮説検定論におよんでくるであろう．これについては，第13章で詳しく議論する．

9.4　標本和の分布の漸近展開

定理 **9.2.1** では，$n\longrightarrow\infty$ とした $(z-n\mu)/\sqrt{n}\,\sigma$ の分布の極限について述べた．ここでは，$(z-n\mu)/\sqrt{n}\,\sigma$ の分布を，十分大きな n に対して，分布 $N(0,1)$ によって与えられる近似より高次まで近似する問題を考えよう．これを p.d.f. を持つ母集団に限定して考察する．

連続な c.d.f. $F(x)$ の中心モーメント $\mu_1,\mu_2,\mu_3,\cdots$ が存在して有限であるとする．$(x-\mu)/\sqrt{n}\,\sigma$ の特性関数を $\varphi(t)$ とすれば

(9.4.1) $$\varphi(t)=1-\frac{t^2}{2n}+\sum_{j=3}^{\infty}\frac{(\mathrm{i}t)^j\alpha_j}{j!(\sqrt{n})^j}.$$

ただし $\alpha_j=\mu_j/\sigma^j$．しかし $(z-n\mu)/\sqrt{n}\,\sigma$ の特性関数 $\varphi_n(t)$ は

(9.4.2) $$\varphi_n(t)=[\varphi(t)]^n$$

で与えられる．対数をとると

(9.4.3) $$\log\varphi_n(t)=-\frac{t^2}{2}+n\sum_{j=3}^{\infty}\frac{(\mathrm{i}t)^j(\kappa_j^*)}{j!(\sqrt{n})^j}.$$

ただし κ_j^* は母集団における $(x-\mu)/\sigma$ の分布の半不変係数である．したがって

(9.4.4) $$\varphi_n(t)=e^{-\frac{1}{2}t^2}\exp\Bigl[n\sum_{j=3}^{\infty}\frac{(\mathrm{i}t)^j(\kappa_j^*)}{j!(\sqrt{n})^j}\Bigr]$$

となり

(9.4.5) $$\varphi_n(t)=e^{-\frac{1}{2}t^2}\Bigl[1+\sum_{j=1}^{\infty}\frac{u_j(\mathrm{i}t)}{(\sqrt{n})^j}\Bigr]$$

と書ける．ここに $u_j(\mathrm{i}t)$ は $(\mathrm{i}t)$ の $3j$ 次多項式で，係数が κ_j^* の関数であるだけで，n には依存しない．$u_j(\mathrm{i}t)$ における $(\mathrm{i}t)$ の最小ベキは $(j+2)$ である．

(9.4.6) $$F_n(x)=P\Bigl(\frac{(z-n\mu)}{\sqrt{n}\,\sigma}\leqslant x\Bigr)$$

と置き，(5.1.14) で $x'=(x+y)/2,\ \delta=(x-y)/2$ とすれば，$x>y$ なる任意の x, y に対して

(9.4.7) $$F_n(x)-F_n(y)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin\left(\dfrac{x-y}{2}\right)t}{t}e^{-\mathrm{i}t[\frac{1}{2}(x+y)]}\varphi_n(t)\,dt$$

を得る．

(9.4.5) を代入して，簡単にすれば

(9.4.8) $$F_n(x)-F_n(y)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{(e^{-\mathrm{i}tx}-e^{-\mathrm{i}ty})}{2(-\mathrm{i}t)}\cdot e^{-\frac{1}{2}t^2}\Bigl[1+\sum_{j=1}^{\infty}\frac{u_j(\mathrm{i}t)}{(\sqrt{n})^j}\Bigr]dt.$$

ここで，$\Phi(x)$ を $N(0,1)$ の c.d.f. とするとき

(9.4.9) $$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{(e^{-\mathrm{i}tx}-e^{-\mathrm{i}ty})}{2(-\mathrm{i}t)}e^{-\frac{1}{2}t^2}\,dt=\Phi(x)-\Phi(y)$$

を示すのは，特に困難ではなかろう．(9.4.7) の他の項は，定数倍をのぞけば

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}(-\mathrm{i}t)^j(e^{-\mathrm{i}tx}-e^{-\mathrm{i}ty})e^{-\frac{1}{2}t^2}\,dt$$

なる型を持つ．これは値

(9.4.10)

$$\frac{1}{2\pi}\frac{d^j}{dx^j}\int_{-\infty}^{\infty}e^{-\mathrm{i}tx-\frac{1}{2}t^2}\,dt-\frac{1}{2\pi}\frac{d^j}{dy^j}\int_{-\infty}^{\infty}e^{-\mathrm{i}ty-\frac{1}{2}t^2}\,dt=\Phi^{(j+1)}(x)-\Phi^{(j+1)}(y)$$

になる．ただし

(9.4.11) $$\Phi^{(j+1)}(x)=\frac{d^{j+1}}{dx^{j+1}}\Phi(x)=\frac{d^j}{dx^j}\Bigl(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\Bigr).$$

したがって $u_j^*(x)$, $u_j^*(y)$ を，$u_j(\mathrm{i}t)$ において $(\mathrm{i}t)^p$ を，それぞれ $\Phi^{(p+1)}(x)$, $\Phi^{(p+1)}(y)$ で置き換えた関数とすれば，(9.4.7) は

(9.4.12) $$F_n(x)-F_n(y)=\Phi(x)-\Phi(y)+\sum_{j=1}^{\infty}\frac{[u_j^*(x)-u_j^*(y)]}{\sqrt{n}^{\,j}}.$$

この式で，$y\longrightarrow-\infty$ とすると，$F_n(x)$ の漸近展開として

(9.4.13) $$F_n(x)=\Phi(x)+\sum_{j=1}^{\infty}\frac{u_j^*(x)}{\sqrt{n}^{\,j}}$$

が得られる．

$F_n(x)$ が p.d.f. $f_n(x)$ を持つならば，(9.4.13) を x で微分すると

$$f_n(x)=\Phi^{(1)}(x)+\sum_{j=1}^{\infty}\frac{u_j^{*\prime}(x)}{\sqrt{n}^{\,j}}. \tag{9.4.14}$$

ただし $u_j^{*\prime}(x)$ は $u_j^*(x)$ の1次微分である．

実際に，$F_n(x)$ が p.d.f. を持たない（すなわち，$F_n(x)$ が離散型 c.d.f. である）場合でも，(9.4.13) は形式的に $F_n(x)$ の連続点で成り立つ．しかし，この場合は，(9.4.14) の展開は意味を持たない．

関数 $u_j^*(x)$ は一般的な表現で書き直すことができないが，(9.4.13)，(9.4.14) の右辺は次数 $n^{-\frac{3}{2}}$ の項まで表わせる．これは

$$\begin{aligned}F_n(x)=\Phi(x)&-\frac{1}{\sqrt{n}}\left(\frac{\alpha_3}{3!}\right)\Phi^{(3)}(x)\\&+\frac{1}{n}\left[\frac{1}{4!}(\alpha_4-3)\Phi^{(4)}(x)+\frac{10}{6!}\alpha_3^2\Phi^{(6)}(x)\right]\\&-\frac{1}{n^{\frac{3}{2}}}\left[\frac{1}{5!}(\alpha_5-10\alpha_3)\Phi^{(5)}(x)+\frac{35}{7!}\alpha_3(\alpha_4-3)\Phi^{(7)}(x)\right.\\&\left.+\frac{280}{9!}\alpha_3^3\Phi^{(9)}(x)\right]+o\left(\frac{1}{n^{\frac{3}{2}}}\right)\cdots\end{aligned} \tag{9.4.15}$$

および

$$\begin{aligned}f_n(x)=\Phi^{(1)}(x)&-\frac{1}{\sqrt{n}}\left(\frac{\alpha_3}{3!}\right)\Phi^{(4)}(x)\\&+\frac{1}{n}\left[\frac{1}{4!}(\alpha_4-3)\Phi^{(5)}(x)+\frac{10}{6!}\alpha_3^2\Phi^{(7)}(x)\right]\\&-\frac{1}{n^{\frac{3}{2}}}\left[\frac{1}{5!}(\alpha_5-10\alpha_3)\Phi^{(6)}(x)+\frac{35}{7!}\alpha_3(\alpha_4-3)\Phi^{(8)}(x)\right.\\&\left.+\frac{280}{9!}\alpha_3^3\Phi^{(10)}(x)\right]+o\left(\frac{1}{n^{\frac{3}{2}}}\right)\cdots.\end{aligned} \tag{9.4.16}$$

Edgeworth (1905) が最初に得たこの結果を要約すると，次を得る．

9.4.1 $(x_1,\cdots,x_n)$ が有限なモーメント $\mu_1,\mu_2,\mu_3,\cdots$ を持つ連続な p.d.f. からの標本のとき，$(z-n\mu)/\sqrt{n}\sigma$ の c.d.f. $F_n(x)$ は (9.4.13) なる型に展開される．また，(9.4.15) なる型で次数 $n^{-\frac{3}{2}}$ の項まで展開することができる．

$\alpha_3\ (=\mu_3/\sigma^3)$ と $\alpha_4-3\ (=\mu_4/\sigma^4-3)$ は，通常，それぞれ γ_1，γ_2 で表わされ，平均

μ，分散 σ^2，3次および4次の中心モーメント μ_3，μ_4 を持つ分布の歪度，尖度と呼ばれている．この2つの定数は，c.d.f. $F_n(x)$ を分布 $N(0,1)$ の c.d.f. $\Phi(x)$ で近似する場合に，重要な役割を果たす．(9.4.15) を考察すると，$F_n(x)$ が次数 $1/\sqrt{n}$ の項をのぞいて，$\Phi(x)$ で近似されることがわかる．しかし，次の系はより高次の近似が成り立つ条件を与えている．

9.4.1 a $(x_1,\cdots,x_n)$ が取り出された分布の歪度が0ならば，$F_n(x)$ は次数 $1/n$ の項をのぞいて，$\Phi(x)$ で近似され，歪度，尖度ともに0ならば，次数 $1/\sqrt{n^3}$ の項をのぞいて，$\Phi(x)$ で近似される．

大標本における $\bar{x}$ の分布を正規分布で高次まで近似する問題は，Lyapunov (1901) が最初に考えた．Cramér (1937) は (9.4.13) および (9.4.14) の残りの項は，無視された第1項と同じ次数であることを示している．Esseen (1944) は漸近展開の精度について研究を続けている．標本平均以外の統計量に関する，$1/\sqrt{n}$ 次までの漸近展開が Cramér (1937)，Hsu (1945 *a*, 1945 *b*)，Chung (1946) により研究されている．Wallace (1958) には漸近展開に関して，幅広く参考文献が集められている．

9.5 有限大母集団からの大標本の線形関数の極限分布

前節では無限母集団からの標本和および標本平均の極限分布について考察した．ここでは，母集団と標本の大きさが限りなく大きくなるとき，標本要素の線形関数の極限分布について考えよう．この問題に関する基本的でしかも一般的な定理は Wald と Wolfowitz (1944) により次のように述べられている．

9.5.1 $(a_{N1},\cdots,a_{NN})$，$(x_{N1},\cdots,x_{NN})$，$N=1,2,\cdots$ は実数からなる2組の集合列で，次の条件を満たすものとする．$r=3,4,\cdots$ および大きな N に対して

$$\frac{m_{r,N}(a)}{[m_{2,N}(a)]^{r/2}}=O(1) \tag{9.5.1}$$

$$\frac{m_{r,N}(x)}{[m_{2,N}(x)]^{r/2}}=O(1). \tag{9.5.2}$$

ただし

(9.5.3) $$m_{r,N}(a)=\frac{1}{N}\sum_{i=1}^{N}(a_{Ni}-\bar{a}_N)^r,\qquad \bar{a}_N=\frac{1}{N}\sum_{i=1}^{N}a_{Ni}$$

で，$m_{r,N}(x)$，$\bar{x}_N$ についても同様に定義されているものとする．各 N に対して，$(x_1,\cdots,x_N)$ は $(x_{N1},\cdots,x_{NN})$ の $N!$ 個の置換全体を標本空間に持つベクトル値確率変数とする．いま

(9.5.4) $$L_N=\sum_{i=1}^{N}a_{Ni}x_i$$

とすれば

(9.5.5) $$\mathscr{E}(L_N)=N\bar{a}_N\bar{x}_N$$

(9.5.6) $$\sigma^2(L_N)=\frac{N^2}{N-1}m_{2,N}(a)\cdot m_{2,N}(x)$$

が成り立つ．さらに

(9.5.7) $$\lim_{N\to\infty}P\left(\frac{L_N-\mathscr{E}(L_N)}{\sigma(L_N)}\leqslant z\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{1}{2}u^2}du$$

である．

9.5.1 の証明は簡単ではあるが，煩雑である．まず，

$$a'_{Ni}=\frac{a_{Ni}-\bar{a}_N}{\sqrt{m_{2,N}(a)}},\qquad x'_{Ni}=\frac{x_{Ni}-\bar{x}_N}{\sqrt{m_{2,N}(x)}}$$

と置くと，実数の集合列 $(a'_{N1},\cdots,a'_{NN})$，$(x'_{N1},\cdots,x'_{NN})$，$N=1,2,\cdots$ も (9.5.1)，(9.5.2) を満たすことが確かめられる．$(x'_1,\cdots,x'_N)$ を $(x_1,\cdots,x_N)$ と同じように，$(x'_{N1},\cdots,x'_{NN})$ の置換を値に持つベクトル値確率変数とする．

$$L'_N=\sum_{i=1}^{N}a'_{Ni}x'_i$$

とすれば，$\mathscr{E}(L'_N)=0$，

(9.5.8) $$\frac{L'_N}{\sigma(L'_N)}=\frac{L_N-\mathscr{E}(L_N)}{\sigma(L_N)}.$$

Wald と Wolfowitz (1944) はこのモーメントを詳しく計算している．それによると，任意の正の整数 k に対して

(9.5.9) $$\mathscr{E}(L'^{s}_N)=\frac{(2k)!}{2^k k!}N^k+o(N^k),\qquad s=2k.$$
$$=o(N^k),\qquad s=2k+1,$$

したがって

(9.5.10) $$\lim_{N\to\infty}\mathscr{E}\left(\frac{L'_N}{\sigma(L'_N)}\right)^s=\lim_{N\to\infty}\mathscr{E}\left(\frac{L_N-\mathscr{E}(L_N)}{\sigma(L_N)}\right)^s=\frac{(2k)!}{2^k k!},\qquad s=2k$$
$$=0,\qquad s=2k+1.$$

このように，$N\longrightarrow\infty$ とすると，$[L_N-\mathscr{E}(L_N)]/\sigma(L_N)$ の s 次のモーメントは分布 $N(0,1)$ を持つ確率変数の s 次のモーメントに収束することがわかる．ゆえに，**5.5.3** より，$n\longrightarrow\infty$ のとき $[L_N-\mathscr{E}(L_N)]/\sigma(L_N)$ の極限分布は $N(0,1)$ になる．

(9.5.4) の L_N の定義で，$a_{Ni}=1/n$，$i=1,\cdots,n$，$a_{Ni}=0$，$i=n+1,\cdots,N$ とすれば，L_N は，$x_{N1},\cdots,x_{NN}$ を x 値に持つ N 個の要素からなる有限母集団 π_N からの大きさ n の標本平均になる．さらに，$\mathscr{E}(L_N)=\mu_N$ (ただし，$\mu_N=\frac{1}{N}\sum_{i=1}^{N}x_{Ni}$)，$\sigma^2(L_N)=(1/n-1/N)\sigma_N^2$ が成立する．ただし，μ_N，σ_N^2 は 8.5 節で定義した π_N の平均，分散である．この場合，(9.5.1) は N，$n\longrightarrow\infty$ のとき，

(9.5.11) $$\frac{\left(\frac{N}{n}-1\right)^{r-1}-(-1)^{r-1}}{\left(\frac{N}{n}\right)\left(\frac{N}{n}-1\right)^{\frac{1}{2}r-1}}$$

の極限値が $r\geqslant 3$ に対して，有限であるという条件に帰着する．$\lim_{N,n\to\infty}\frac{N}{n}=c$，$+1<c<+\infty$ ならば，この条件は満たされる．したがって，**9.5.1** の系として，次を得る．

9.5.1 a　$\{\pi_N,\ N=1,2,\cdots\}$ は平均 μ_N，分散 σ_N^2 を持つ有限母集団の列とし，$\bar{x}_n$ は π_N からの大きさ n の標本平均とする．このとき，$(x_{N1},\cdots,x_{NN},\ N=1,2,\cdots)$ が，十分大きな N に対して (9.5.2) を満たし，$\lim_{N,n}N/n=c$，$1<c<+\infty$ となるように N，$n\longrightarrow\infty$ とすれば，

(9.5.12) $$\lim P\left(\frac{(\bar{x}_n-\mu_N)}{\left[\left(\frac{1}{n}-\frac{1}{N}\right)\sigma_N^2\right]^{\frac{1}{2}}}\leqslant z\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{1}{2}u^2}du$$

を満たす．

9.6　順序統計量の漸近分布

(a)　カバーの和の極限分布

8.7.2 では，連続な c.d.f. $F(x)$ からの順序統計量を $(x_{(1)},\cdots,x_{(n)})$ としたとき，確率

変数 $y_{(k)}=F(x_{(k)})$ は c.d.f.

(9.6.1)
$$D_n(y_{(k)})=\frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)}\int_0^{y(k)}x^{k-1}(1-x)^{n-k}\,dx$$

を持つことが示された．事実，**8.7.6** に着目すると，(9.6.1) 式は，$(n+1)$ 個のカバー $F(x_{(1)}),F(x_{(2)})-F(x_{(1)}),\cdots,1-F(x_{(n)})$ の任意の k 個の和の c.d.f. であることがわかる．$0<s<n-k$ とすれば，$F(x_{(s+k)})-F(x_{(s)})$ がこのような和として考えられる．さて，確率変数 $w_k=ny_{(k)}$ の c.d.f. を $H_n(w_k)$ で表わせば

(9.6.2)
$$H_n(w_k)=D_n\left(\frac{w_k}{n}\right)=\int_0^{w_k}h_n(y)\,dy$$

となる．ただし

$$h_n(y)=\frac{\Gamma(n+1)}{n^k\Gamma(k)\Gamma(n-k+1)}y^{k-1}\left(1-\frac{y}{n}\right)^{n-k}.$$

$n\longrightarrow\infty$ とすれば，列 $h_n(y)$, $n=k,k+1,\cdots$ は区間 $(0,w_k)$ 上で関数 $\dfrac{1}{\Gamma(k)}y^{k-1}e^{-y}$ に一様収束することは明らかである．よって

(9.6.3)
$$\lim_{n\to\infty}H_n(w_k)=\frac{1}{\Gamma(k)}\int_0^{w_k}y^{k-1}e^{-y}\,dy.$$

したがって，次の結果を得る．

9.6.1　$(x_{(1)},\cdots,x_{(n)})$ が連続な c.d.f. $F(x)$ を持つ母集団からの標本順序統計量とすれば，固定された k に対して，$nF(x_{(k)})$, $n=k,k+1,\cdots$ はガンマ分布 $G(k)$ に法則収束する確率変数列になる．

9.6.1 は，$F(x_{(k)})$ のかわりに，カバー $F(x_{(1)}),F(x_{(2)})-F(x_{(1)}),\cdots,1-F(x_{(n)})$ の任意の k 個の和，たとえば，$1-F(x_{(n-k+1)})$ または $F(x_{(s+k)})-F(x_{(s)})$ などに置き換えても成り立つ．また，連続な多次元 c.d.f. を持つ母集団からの大きさ n の標本より定まる任意の k 個のカバーに換えても，**9.6.1** がいえる [8.7(c) 節を参照せよ]．

定理 **9.6.1** はカバーの固定された個数の和が2つ以上の場合に拡張される．2つの場合については，次のように同様な結果を得る．

9.6.2　$(x_{(1)},\cdots,x_{(n)})$ は連続な c.d.f. $F(x)$ からの標本順序統計量とする．k_1, k_2 を固定した整数とすれば，2次元確率変数列 $(nF(x_{(k_1)}),nF(x_{(k_1+k_2)}))$, $n=m,m+1,\cdots$ $(m\geqslant k_1+k_2)$ は，次の p.d.f. $f(w_1,w_2)$ を持つ分布に法則収束する．

$$f(w_1,w_2)=\begin{cases}\dfrac{1}{\Gamma(k_1)\Gamma(k_2)}w_1^{k_1-1}(w_2-w_1)^{k_2-1}e^{-w_2}, & 0<w_1<w_2<\infty\\ 0, & \text{その他}.\end{cases}$$

この証明は **9.6.1** で用いた方法と同様である．読者はこれを確かめよ．**9.6.2** では，特に，$F(x_{(k_1)})$ を任意の k_1 個のカバーの和に，$F(x_{(k_1+k_2)})$ を最初の k_1 個のカバーを含む任意の (k_1+k_2) 個のカバーの和に置き換えても成立することに注意せよ．

さて，k を固定しないで $k/n=p_n=p+O\left(\frac{1}{n}\right)$ となるように，k と n を増加させるときの $y_{(k)}$ の極限分布（ただし $n\longrightarrow\infty$）を考察しよう．確率変数列 $v_n=(y_{(np_n)}-p)\sqrt{n}$, $n=m,m+1,\cdots;m\geqslant np_n$ を考える．ただし，$y_{(np_n)}=F(x_{(np_n)})$. v_n の c.d.f. を $H_n^*(v)$ とすれば

(9.6.4)
$$H_n^*(v)=D_n\left(p+\frac{v}{\sqrt{n}}\right).$$

ただし，$D_n(y_{(np_n)})$ は (9.6.1) で定義されている．$D_n(p+v/\sqrt{n})$ は

(9.6.5)
$$\int_{-p\sqrt{n}}^{v}h_n^*(z)\,dz$$

と書ける．ここに

(9.6.6)
$$\begin{aligned}h_n^*(z)&=\frac{\Gamma(n+1)}{\sqrt{n}\,\Gamma(k)\Gamma(n-k+1)}\left(p+\frac{z}{\sqrt{n}}\right)^{k-1}\left(1-p-\frac{z}{\sqrt{n}}\right)^{n-k}\\&=\frac{\Gamma(n+1)(p+z/\sqrt{n})^{-1}}{\sqrt{n}\,\Gamma(p_nn)\Gamma((1-p_n)n+1)}[p^{p_n}(1-p)^{1-p_n}]^n\\&\quad\cdot\left[\left(1+\frac{z}{p\sqrt{n}}\right)^{p_n}\left(1-\frac{z}{(1-p)\sqrt{n}}\right)^{1-p_n}\right]^n.\end{aligned}$$

しかし，$p_n=p+O\left(\frac{1}{n}\right)$ だから

(9.6.7)
$$\left(1+\frac{z}{p\sqrt{n}}\right)^{p_n}\left(1-\frac{z}{(1-p)\sqrt{n}}\right)^{1-p_n}=\left(1-\frac{z^2}{2np(1-p)}+\varphi(z,n)\right).$$

ただし $\varphi(z,n)$ は任意の z に対し $\lim\limits_{n\to\infty}n\varphi(z,n)=0$ である．g が大きいとき，$\Gamma(g)$ に対するスターリングの近似式 (7.6.27) を用いれば

(9.6.8)
$$\frac{\Gamma(n+1)}{\sqrt{n}\,\Gamma(p_nn)\Gamma((1-p_n)n+1)}[p^{p_n}(1-p)^{1-p_n}]^n=\sqrt{\frac{p}{2\pi(1-p)}}+O\left(\frac{1}{n}\right).$$

ゆえに

(9.6.9)

$$h_n^*(z) = \left(p + \frac{z}{\sqrt{n}}\right)^{-1}\left(\sqrt{\frac{p}{2\pi(1-p)}} + O\left(\frac{1}{n}\right)\right)\left(1 - \frac{z^2}{2np(1-p)} + \varphi(z, n)\right)^n$$

を得る．これは $n \longrightarrow \infty$ とした任意の区間 $(-K, v)$ で関数

(9.6.10)
$$h^*(z) = \frac{1}{\sqrt{2\pi p(1-p)}} e^{-z^2/2p(1-p)}$$

に一様収束する．任意の $\varepsilon > 0$ に対して，K と n_1 を選べば，$n > n_1$ に対して，

(9.6.11)
$$\int_{-\infty}^{-K} h^*(z) dz < \frac{\varepsilon}{3}, \qquad \left|\int_{-K}^{v} h_n^*(z)\, dz - \int_{-p\sqrt{n}}^{v} h_n^*(z)\, dz\right| < \frac{\varepsilon}{3}$$

が成り立つようにできる．さらに $n_2 > K^2/p^2$ が存在して任意の v，$n > n_2$ に対して

(9.6.12)
$$\left|\int_{-K}^{v} h_n^*(z)\, dz - \int_{-K}^{v} h^*(z)\, dz\right| < \frac{\varepsilon}{3}$$

となる．ゆえに，$n > \max(n_1, n_2)$ ならば

(9.6.13)
$$\left|\int_{-\infty}^{v} h^*(z)\, dz - \int_{-p\sqrt{n}}^{v} h_n^*(z)\, dz\right| < \varepsilon.$$

よって

(9.6.14)
$$\lim_{n\to\infty} P\left(\frac{[F(x_{(np_n)}) - p]\sqrt{n}}{\sqrt{p(1-p)}} < w\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{w} e^{-\frac{1}{2}u^2}\, du$$

である．要約すると

9.6.3　$(x_{(1)}, \cdots, x_{(n)})$ は連続な c.d.f. $F(x)$ からの標本順序統計量とする．$p_n = p + O\left(\frac{1}{n}\right)$ $(0 < p < 1)$ となるようにして，しかも np_n が整数であるようにする．このとき，大きな n に対し，$F(x_{(np_n)})$ は漸近的に $N\left(p, \frac{p(1-p)}{n}\right)$ に従う．

9.6.3 は1次元および多次元においても，あるいは，$F(x_{(np_n)})$ を任意の np_n 個のカバーの和にしても成立することに注意せよ．

9.6.3 はカバーの和の2つ以上からなる組に対しても容易に拡張される．2つの場合を述べれば十分であろう．読者はこれを証明せよ．

9.6.4　$(x_{(1)}, \cdots, x_{(n)})$ は連続な c.d.f. $F(x)$ を持つ母集団からの標本順序統計量とする．np_{1n}，np_{2n} はそれぞれ $p_{1n} = p_1 + O\left(\frac{1}{n}\right)$，$p_{2n} = p_2 + O\left(\frac{1}{n}\right)$ $(0 < p_1 < p_2 < 1)$ となる整数値とする．このとき，大きな n に対して，2次元確率変数 $(F(x_{(np_{1n})}), F(x_{(np_{2n})}))$ は漸近的分布 $N\left(\{p_i\}; \left\|\frac{\sigma_{ij}}{n}\right\|\right)$ を持つ．ただし $\sigma_{11} = p_1(1-p_1)$，$\sigma_{12} = \sigma_{21} = p_1(1-p_2)$，$\sigma_{22} = p_2(1-p_2)$．

(b)　順序統計量の極限分布

9.6.1, **9.6.2**, **9.6.3**, **9.6.4** で得られた極限分布は，直接には，カバーの和の分布に関する性質である．しかし，暗に，大標本順序統計量に関する性質も示されていた．母集団分布 $F(x)$ の連続性を仮定しているので，順序統計量の極限分布に関しても述べることができる．たとえば，**9.6.1** で $F(x)$ の逆関数を $F^{-1}(y)$ とすれば，$F^{-1}(k/n)$ が唯一の逆元を持つ n の列を考えることができる．$F(x)$ は連続ですべての実数 x に対して定義されているから，このような n の無限列，たとえば，$n_1, n_2, \cdots$ が存在する．よって，**9.6.1** から固定された k に対して

(9.6.15)
$$\lim_{i\to\infty} P\left(x_{(k)} \leqslant F^{-1}\left(\frac{u}{n_i}\right)\right) = \frac{1}{\Gamma(k)} \int_0^u y^{k-1} e^{-y}\, dy$$

が成り立つ．ゆえに，大きな n に対して

(9.6.16)
$$P(x_{(k)} \leqslant v) \cong \frac{1}{\Gamma(k)} \int_0^{nF(v)} y^{k-1} e^{-y}\, dy$$

を得る．k' を固定すると，$x_{(n-k'+1)}$ に対しても，(9.6.15) および (9.6.16) と同様な式が得られる．このような問題に関して漸近的な結果，特に，$k = 1$（最小順序統計量），$k' = 1$（最大順序統計量）の場合などが，Dodd (1923)，Fisher と Tippett (1928)，Fréchet (1927)，Gumbel (1935, 1958)，Smirnov (1935) により，詳しく研究されている．

同じように，**9.6.3** から

(9.6.17)
$$\lim_{i\to\infty} P\left(x_{(n_i p_{n_i})} \leqslant F^{-1}\left(p + \frac{w\sqrt{p(1-p)}}{\sqrt{n_i}}\right)\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{w} e^{-\frac{1}{2}y^2}\, dy$$

を得る．また，大きな n に対して

(9.6.18)
$$P(x_{(np_n)} \leqslant v) \cong \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{T} e^{-\frac{1}{2}y^2}\, dy$$

ただし

$$T = \frac{(F(v) - p)\sqrt{n}}{\sqrt{p(1-p)}}.$$

公式 (9.6.16)，(9.6.18) を2つ以上の順序統計量に拡張することは難しくない．2次元の場合でも，**9.6.2**, **9.6.4** からすぐ成り立つ．

(9.6.14) より，$n \longrightarrow \infty$ のとき，$F(x_{(np_n)})$ は定数 p に確率収束する．ここで $F(x) = p$ が一意の解を持つ，すなわち，p 分位点 $\underline{x}_p$ が一意に存在すると仮定しよう．$F(x)$ は連続だから，$x_{(np_n)}$ は $\underline{x}_p$ に確率収束する．すなわち，

9.6.5　　**9.6.3** の仮定の下に，p 分位点 $\underline{x}_p$ が一意に存在すれば，$n \longrightarrow \infty$ のとき，$x_{(np_n)}$ は $\underline{x}_p$ に確率収束する．

同様の結果が 2 つの順序統計量 $(x_{(np_{1n})}, x_{(np_{2n})})$ についても成り立つ．3 つ以上についても同様である．

さて，$F(x)$ は $x = \underline{x}_p$ のある近傍 $V(\underline{x}_p)$ で導関数 $f(x)$ を持ち，$f(\underline{x}_p) > 0$ とする．唯一の p 分位点 $\underline{x}_p$ が存在するから，**9.6.5** により，$x_{(np_n)}$ は $\underline{x}_p$ に確率収束する．$x_{(np_n)}$ が $V(\underline{x}_p)$ の任意の点であれば

(9.6.19)　　$$F(x_{(np_n)}) = p + f(x^*)(x_{(np_n)} - \underline{x}_p)$$

と書ける．ただし x^* は $|x^* - \underline{x}_p| < |x_{(np_n)} - \underline{x}_p|$ となる確率変数である．しかし $x_{(np_n)}$ は $n \longrightarrow \infty$ のとき p 分位点 $\underline{x}_p$ に確率収束するので，任意の $\varepsilon > 0$ に対して，n_ε を選んで

(9.6.20)　　$$P(F(x_{(np_n)}) = p + f(x^*)(x_{(np_n)} - \underline{x}_p) \text{ すべての } n > n_\varepsilon) > 1 - \varepsilon$$

となるようにできる．このことは

(9.6.21)　　$$\frac{(F(x_{(np_n)}) - p)\sqrt{n}}{\sqrt{p(1-p)}}, \qquad n = m, m+1, \cdots, \qquad m \geqslant np_n$$

(9.6.22)　　$$\frac{f(x^*)(x_{(np_n)} - \underline{x}_p)\sqrt{n}}{\sqrt{p(1-p)}}, \qquad n = m, m+1, \cdots, \qquad m \geqslant np_n$$

がともに，分布 $N(0,1)$ に法則収束する確率過程を構成していることを示している．

$F(x)$ は $V(\underline{x}_p)$ で導関数 $f(x)$ を持っているので，$f(x)$ は $V(\underline{x}_p)$ で，したがって $x = \underline{x}_p$ で，連続である．よって，**4.3.5** により，$f(x_{(np_n)})$ は定数 $f(\underline{x}_p)$ に確率収束する．**4.3.7** から，$f(x^*)$ もまた $n \longrightarrow \infty$ のとき，$f(\underline{x}_p)$ に確率収束する．ゆえに，**4.3.3** を適用すれば，確率変数列 (9.6.22) と列

(9.6.22 a)　　$$\frac{f(\underline{x}_p)(x_{(np_n)} - \underline{x}_p)\sqrt{n}}{\sqrt{p(1-p)}}, \qquad n = m, m+1, \cdots, \qquad m \geqslant np_n$$

は，ともに分布 $N(0,1)$ に法則収束する確率過程を構成することとなる．

要約すると，次のようになる．

9.6.6　　**9.6.3** の仮定に加えて，$F(x)$ が $x = \underline{x}_p$ のある近傍 $V(\underline{x}_p)$ で，$f(\underline{x}_p) > 0$ となる導関数 $f(x)$ を持つとすれば，十分大きな n に対して，$x_{(np_n)}$ は漸近的に分布 $N\left(\underline{x}_p, \dfrac{p(1-p)}{nf^2(\underline{x}_p)}\right)$ に従う．

例題　$p = \dfrac{1}{2}$ のとき，**9.6.6** の興味ある結果が得られる．この場合，$x_{(np_n)}$ は標本メディアンであり，大きな n に対し漸近的分布 $N\left(\underline{x}_{0.5}, \dfrac{1}{4nf^2(\underline{x}_{0.5})}\right)$ を持つことがわかる．母集団分布が $N(\mu, \sigma^2)$ ならば，標本メディアンは，大きな n に対して漸近的分布 $N(\mu, \pi\sigma^2/2n)$ を持つ．

もちろん，**9.6.6** は 2 つ以上の順序統計量の結合分布に関しても成り立つ．2 つの順序統計量 $x_{(np_{1n})}$，$x_{(np_{2n})}$ の場合，次のように述べることができる．

9.6.7　　**9.6.4** の仮定に加えて，$F(x)$ が $x = \underline{x}_{p_1}$，$x = \underline{x}_{p_2}$ $(0 < p_1 < p_2 < 1)$ のそれぞれの近傍で導関数 $f(x)$ を持ち，$f(\underline{x}_{p_1}) > 0$，$f(\underline{x}_{p_2}) > 0$ とする．このとき，十分大きな n に対し，確率変数 $(x_{(np_{1n})}, x_{(np_{2n})})$ は漸近的に $N\left(\{\underline{x}_{p_i}\}, \left\|\dfrac{\sigma^*_{ij}}{x}\right\|\right)$，$i, j = 1, 2$ に従う分布を持つ．ただし

$$\sigma^*_{11} = \frac{p_1(1-p_1)}{f^2(\underline{x}_{p_1})}, \qquad \sigma^*_{12} = \frac{p_1(1-p_2)}{f(\underline{x}_{p_1})f(\underline{x}_{p_2})}, \qquad \sigma^*_{22} = \frac{p_2(1-p_2)}{f^2(\underline{x}_{p_2})}.$$

定理 **9.6.6** と **9.6.7** は Smirnov (1935) によって得られた．しかし，$x_{(np_{1n})}$ と $x_{(np_{2n})}$ の分散および共分散については，K. Pearson (1920) が最初に明らかにした．Mosteller (1946) は **9.6.7** を 3 つ以上の順序統計量の場合に拡張している．

問　　題

9.1　**9.1.1** を k 次元の場合について述べて証明せよ．

9.2　$(x_{1\xi}, \cdots, x_{k\xi}; \xi = 1, \cdots, n)$ は，有限な平均 $(\mu_1, \cdots, \mu_k)$，有限な共分散行列 $\|\sigma_{ij}\|$ を持つ k 次元分布からの標本とする．また $(\bar{x}_1, \cdots, \bar{x}_k)$ は標本平均ベクトルである．このとき，任意の $\delta_i(>0)$，$i = 1, \cdots, k$ に対し

$$P(|\bar{x}_i - \mu_i| < \delta_i,\ i = 1, \cdots, k) \geqslant 1 - \frac{1}{n}\left(\frac{\sigma_{11}}{\delta_1^2} + \cdots + \frac{\sigma_{kk}}{\delta_k^2}\right).$$

したがって，$(\bar{x}_1, \cdots, \bar{x}_k)$ が $(\mu_1, \cdots, \mu_k)$ に確率収束することを示せ．

9.3（続き）　任意の $\delta^2 > 0$ に対して

$$P\left(\sum_{i,j=1}^{k} \sigma^{ij}(\bar{x}_i - \mu_i)(\bar{x}_j - \mu_j) < \delta^2\right) \geqslant 1 - \frac{k}{n\delta^2}.$$

したがって，$(\bar{x}_1, \cdots, \bar{x}_k)$ が $(\mu_1, \cdots, \mu_k)$ に確率収束することを示せ．ただし $\|\sigma^{ij}\| =$

$\|\sigma_{ij}\|^{-1}$.

9.4　**9.2.3** を証明せよ.

9.5　$(x_1,\cdots,x_n)$ がポアソン分布 $Po(\mu)$ からの標本のとき，大きな n に対し，$2\sqrt{\bar{x}}$ は漸近的分布として $N\left(2\sqrt{\mu},\dfrac{1}{n}\right)$ を持つことを示せ．$(x_1,\cdots,x_n)$ がガンマ分布 $G(\mu)$ からの標本のときも，同じ結果が成り立つことを示せ.

9.6　$(x_1,\cdots,x_n)$ が2項分布 $Bi(1,p)$ からの標本ならば，$\sin^{-1}(2\bar{x}-1)$ の漸近分布は，大きな n に対し，$N\left(\sin^{-1}(2p-1),\dfrac{1}{n}\right)$ であることを示せ.

9.7　$(x_1,\cdots,x_n)$ が p.f. qp^{x-1}, $x=1,2,\cdots$ を持つ待ち時間分布からの標本のとき，n が大きいならば $\log\left[\bar{x}(1+\sqrt{1-1/\bar{x}})-\dfrac{1}{2}\right]$ の漸近的分布は

$$N\left(\log\left[\frac{1+\sqrt{p}}{q}-\frac{1}{2}\right],\frac{1}{n}\right)$$

であることを示せ．ただし $0<p<1$, $p+q=1$.

9.8　$(x_1,\cdots,x_n)$ が矩形分布 $R\left(\dfrac{1}{2}\theta,\theta\right)$ からの標本ならば，十分大きな n に対して $\sqrt{12}\log(2\bar{x})$ の漸近的分布は $N\left(\sqrt{12}\log\theta,\dfrac{4}{n}\right)$ であることを示せ.

9.9　$\bar{x}$ が矩形分布 $R\left(\dfrac{1}{2},1\right)$ からの大きさ n の標本の標本平均のとき，確率変数 $\left(\bar{x}-\dfrac{1}{2}\right)\sqrt{12n}$ の c.d.f. と p.d.f. に対して (9.4.15), (9.4.16) の展開を $n^{-\frac{3}{2}}$ の次数まで求めよ.

9.10　$(x_1,\cdots,x_n)$ が有限な平均値 μ, 分散 σ^2 を持つ分布からの標本ならば，標本分散 s^2 は σ^2 に確率収束することを示せ．また，大きな n に対し，$\sqrt{n}(\bar{x}-\mu)/s$ は漸近分布として $N(0,1)$ を持つことを示せ.

9.11　n_1, $\bar{x}_1$, s_1^2 は平均 μ_1, 分散 σ_1^2 を持つ分布からの，それぞれ，標本の大きさ，標本平均，標本分散とする．一方，n_2, $\bar{x}_2$, s_2^2 は平均 μ_2, 分散 σ_2^2 を持つ分布からの，前述の標本と独立な標本の大きさ，平均，分散とする．$n_1,n_2\longrightarrow\infty$ のとき

$$\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}$$

の極限分布は $N(0,1)$ であることを示せ.

9.12　$\tilde{x}$ が連続な c.d.f. $F(x)$ からの大きさ n の標本メディアンならば，大きな n に対し，$F(\tilde{x})$ の漸近分布は $N\left(\dfrac{1}{2},\dfrac{1}{(4n)}\right)$ であることを示せ.

9.13　$(x_{(1)},\cdots,x_{(n)})$ は連続な c.d.f. $F(x)$ からの標本順序統計量とする．$n\longrightarrow\infty$ のとき，$n\left(1-\int_{x_{(1)}}^{x_{(n)}}dF(x)\right)$ の極限分布はガンマ分布 $G(2)$ であることを示せ.

9.14　$(x_1,\cdots,x_{2n+1})$ は p.d.f. $\lambda e^{-\lambda x}$ $(x>0,\ \lambda>0)$ を持つ分布からの標本とする．このとき，大きな n に対し，標本メディアン $\tilde{x}$ は漸近的分布として，$N(\log 2/\lambda, 1/(2\lambda^2 n))$ を持つことを示せ.

9.15　$\bar{x}$, $\tilde{x}$ は分布 $N(\mu,\sigma^2)$ を持つ母集団からの大きさ n の標本平均，標本メディアンとする．大きな n に対し，$(\bar{x},\tilde{x})$ の漸近分布は

$$N\left(\mu,\mu;\left\|\frac{\sigma_{ij}}{n}\right\|\right)$$

であることを示せ．ただし

$$\left\|\frac{\sigma_{ij}}{n}\right\|=\left\|\begin{matrix}\dfrac{\sigma^2}{n} & \dfrac{\sigma^2}{n}\\[2mm] \dfrac{\sigma^2}{n} & \dfrac{\pi\sigma^2}{2n}\end{matrix}\right\|.$$

9.16　**(相関係数の Fisher (1925 *a*) 変換)**　相関係数 ρ を持つ2次元正規分布からの標本相関係数 r の分布は，大きな n のとき，分布 $N(\rho,(1-\rho^2)^2/n)$ を漸近分布として持つことが知られている．大きな n に対し，$\dfrac{1}{2}\log\left(\dfrac{1+r}{1-r}\right)$ の漸近分布は

$$N\left(\frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right),\frac{1}{n}\right)$$

になることを示せ.

9.17　**9.6.2** を証明せよ.

9.18　**9.6.4** を証明せよ.

参考文献

[　]の中の数字は本文の引用頁を表わす.

Army Ordnance Corps (1952), *Tables of the Cumulative Binomial Probabilities*, ORDP, 20-11. [139]

I. L. Battin (1942), On the problem of multiple matching, *Ann. Math. Stat.*, Vol. 13, pp. 294-305. [154]

J. Bernoulli (1713), *Ars Conjectandi*, Paris. [138]

G. Birkhoff and S. MacLane (1953), *A Survey of Modern Algebra* (Revised Edition), Macmillan, New York. [210]

M. Bôcher (1907), *Introduction to Higher Algebra*, Macmillan, New York. [91, 210]

R. C. Bose (1938), On the application of the properties of Galois fields to the problem of construction of hyper-Graeco-Latin squares, *Sankyhā*, Vol. 3, pp. 323-338. [230]

R. C. Bose (1939), On the construction of balanced incomplete block designs, *Ann. Eugen.*, Vol. 9, pp. 353-399. [229]

G. E. P. Box and M. E. Muller (1958), A note on the generation of random normal deviates, *Ann. Math. Stat.*, Vol. 29, pp. 610-611. [187]

C. Carathéodory (1927). *Vorlesungen über reelle Funktionen*, Leipzig-Berlin. [15]

A. G. Carlton (1946), Estimating the parameters of a rectangular distribution, *Ann. Math. Stat.*, Vol. 17, pp. 355-358. [244]

A. L. Cauchy (1853), Sur les résultats moyens d'observations de méme nature et sur les résultats les plus probables, *Comp. Rend. Acad. Sci.*, Paris, Vol. 37, pp. 198-206. [130, 251]

P. L. Chebyshev (1867), Des valeurs moyennes, *Jour. Math. Pures et Appl.*, Vol. 12, pp. 177-184. [76]

P. L. Chebyshev (1890), Sur deux théorèmes relatifs aux probabilités. *Acta Math.*, Vol. 14, pp. 305-315. [253]

K. L. Chung (1941), On the probability of the occurrence of at least m events among n arbitrary events, *Ann. Math. Stat.*, Vol. 17, pp. 447-465. [30]

K. L. Chung (1946), The approximate distribution of Student's statistic, *Ann. Math. Stat.*, Vol. 17, pp. 447-465. [261]

W. G. Cochran (1934), The distribution of quadratic forms in a normal system, with applications to the analysis of covariance, *Proc. Camb. Phil. Soc.*, Vol. 30, pp. 178-191. [209]

W. G. Cochran and G. M. Cox (1957), *Experimental Designs*, second edition, John Wiley, New York. [229]

W. S. Connor, Jr. (1952), On the structure of balanced incomplete block designs, *Ann. Mash. Stat.*, Vol. 23, pp. 57-71. [229]

E. A. Cornish and R. A. Fisher (1937), Moments and cumulants in the specification of distributions, *Rev. Int. Stat. Inst.*, Vol. 4, pp. 1-14. [198]

A. T. Craig (1932), On the distributions of certain statistics, *Amer. Jour. Math.*, Vol. 54, pp. 353-366. [233]

C. C. Craig (1928), An application of Thiele's semivariants to the sampling problem, *Metron*, Vol. 7, No. 4, pp. 3-74. [198]

H. Cramér (1936), Über eine Eigenschaft der normalen Verteilungsfunktion, *Math. Zeits.*, Vol. 41, pp. 405-414. [208]

H. Cramér (1937), *Random Variables and Probability Distributions*, Cambridge Tracts in Mathematics, No. 36, Cambridge University Press. [123, 261]

H. Cramér (1946), *Mathematical Methods of Statistics*, Princeton University Press. [93, 94, 124, 126, 132, 176]

A. DeMoivre (1718), *The Doctrine of Chances*, London. [253]

P. G. L. Dirichlet (1839), Sur un nouvelle methode pour la determination des integrales multiples, *Comp. Rend. Acad. Sci.*, Vol. 8, pp. 156-160. [178]

E. L. Dodd (1923), The greatest and least variate under general laws of error, *Trans. Amer. Math. Soc.*, Vol. 25, pp. 525-539. [267]

J. L. Doob (1953), *Stochastic Processes*, John Wiley, New York. [1, 26, 99, 100]

R. Dorfman (1943), The detection of defective members of large populations, *Ann. Math. Stat.*, Vol. 14, pp. 436-440. [152]

P. S. Dwyer (1938), Combined expansions of products of symmetric power sums and of sums of symmetric power products with applications to sampling, *Ann. Math. Stat.*, Vol. 9, pp. 1-47, 97-132. [198]

F. Y. Edgeworth (1905), The law of error, *Trans. Camb. Phil. Soc.*, Vol. 20, pp. 36-65. [260]

G. Elfving (1947), The asymptotical distribution of range in samples from a normal population, *Biometrika*, Vol. 34, pp. 111-119. [244]

B. Epstein (1954), Tables for the distribution of the number of exceedances, *Ann. Math. Stat.*, Vol. 25, pp. 762-768. [247]

C. E. Esseen (1944), Fourier analysis of distribution functions: A mathematical study of the Laplace-Gaussian law, *Acta. Math.*, Vol. 77, pp. 1-125. [261]

W. Feller (1935), Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung, *Math. Zeits.*, Vol. 40, pp. 521-559. [253, 254]

W. Feller (1957), *An Introduction to Probability Theory and its Applications*, second edition, John Wiley, New York. [1, 100, 110, 192]

邦訳 国沢清典監訳 現代経営科学全集《確率論とその応用 (II) 下》第1版 紀伊國屋

R. A. Fisher (1915), Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population, *Biometrika*, Vol. 10, pp. 507-521. [193]

R. A. Fisher (1924), On a distribution yielding the error functions of several well-known statistics, *Proc. Int. Math. Congress*, Vol. II, Toronto, pp. 805-813. [186]

R. A. Fisher (1925 *a*), *Statistical Methods for Research Workers*, first edition, twelfth edition (1954), Oliver and Boyd, Edinburgh. [182, 184, 186, 271]

邦訳 遠藤健児・鍋谷清治共訳《研究者のための統計的方法》森北出版

R. A. Fisher (1926 *a*), Applications of "Student's" distribution, *Metron*, Vol. 5, No. 4, pp. 90-104. [208]

R. A. Fisher (1928 *a*), Moments and product moments of sampling distributions, *Proc. London Math. Soc.*, Vol. 30, pp. 199-238. [198]

R. A. Fisher (1928 *b*), The general sampling distribution of the multiple correlation coefficient, *Proc. Roy. Soc. London, Series A*, Vol. 121, pp. 654-673. [243]

R. A. Fisher and L. H. C. Tippett (1928), Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Camb. Phil. Soc.*, Vol. 24, pp. 180-190. [267]

E. C. Fieller (1932), The distribution of the index in a normal bivariate population, *Biometrika*, Vol. 24, pp. 428-440. [187]

D. A. S. Fraser (1951), Sequentially determined statistically equivalent blocks, *Ann. Math. Stat.*, Vol. 22, pp. 372-381. [239]

D. A. S. Fraser (1953), Nonparametric tolerance regions, *Ann. Math. Stat.*, Vol. 24, pp. 44-55. [239]

D. A. S. Fraser and I. Guttman (1956), Tolerance regions, *Ann. Math. Stat.*, Vol. 27, pp. 162-179. [239]

D. A. S. Fraser (1957), *Nonparametric Methods in Statistics*, John Wiley, New York. [231]

M. Fréchet (1927), Sur la loi de probabilitié de l'écart maximum, *Ann. Soc. Polonaise Math.*, Vol. 6, pp. 92-116. [267]

K. F. Gauss (1809 *b*), *Werke*, Vol. 4, Göttingen, pp. 1-93. [253]

R. C. Geary (1936), Distribution of Student's distribution for non-normal samples, *Jour. Roy. Stat. Soc., Series B*, Vol. 3, pp. 178-184. [208]

M. A. Girshick, F. Mosteller, and L. J. Savage (1946), Unbiased estimates for certain binomial sampling problems with applications. *Ann. Math. Stat.*, Vol. 17, pp. 13-23. [145]

B. V. Gnedenko and A. N. Kolmogorov (1954), *Limit Distributions for Sums of*

Independent Random Variables (translated from the 1949 Russian edition by K. L. Chung; with an appendix by J. L. Doob), Addison-Wesley Publishing Company, Reading, Mass. [1, 188, 253]

W. S. Gosset (1908), "Student," The probable error of a mean, *Biometrika*, Vol. 6, pp. 1-25. [208]

J. A. Greenwood and H. O. Hartley (1961), *Guide to Tables in Mathematical Statistics*, Princeton University Press. [157, 186]

E. J. Gumbel (1935), Les valeurs extrêmes des distributions statistiques, *Ann. de l'Institut Henri Poincaré*, Vol. 4, pp. 115-158. [267]

E. J. Gumbel (1958), *Statistics of Extremes*, Columbia University Press. [231, 267]

I. Guttman (1960), Personal communication. [245]

J. B. S. Haldane (1945), On a method of estimating frequencies, *Biometrika*, Vol. 33, pp. 222-225. [145]

P. R. Halmos (1950), *Measure Theory*, D. Van Nostrand, Princeton, N. J. [1, 15, 16, 21, 26]

T. E. Harris (1948), Branching processes, *Ann. Math. Stat.*, Vol. 19, pp. 474-494. [133]

H. O. Hartley and H. A. David (1954), Universal bounds for mean range and extreme observation, *Ann. Math. Stat.*, Vol. 25, pp. 85-99. [246]

Harvard Computation Laboratory (1955), *Tables of the Cumulative Binomial Probability Distribution*, Harvard University. [139]

F. R. Helmert (1876 *a*), Ueber die Wahrscheinlichkeit der Potensummen der Beobachtungsfehler und über einige damit im Zusammenhange stehende Fragen, *Zeits. für Math. und Phys.*, Vol. 21, pp. 192-218. [206]

F. R. Helmert (1876 *b*), Die Genauigkeit der Formel von Peters zur Berechnung des Wahrscheinlichen Beobachtungsfehlers directer Beobachtungen glicher Genauigkeit, *Astron. Nachr.*, Vol. 88, pp. 112-131. [208]

W. Hoeffding (1948 *a*), A class of statistics with asymptotically normal distribution, *Ann. Math. Stat.*, Vol. 19, pp. 293-325. [200]

R. Hooke (1956 *a*), Symmetric functions of a two-way array, *Ann. Math. Stat.*, Vol. 27, pp. 55-79. [223]

P. L. Hsu (1945 *a*), The approximate distributions of the mean and variance of a sample of independent variables, *Ann. Math. Stat.*, Vol. 16, pp. 1-29. [261]

P. L. Hsu (1945 *b*), The asymptotic distribution of ratios, *Ann. Math. Stat.*, Vol. 16, pp. 204-210. [261]

J. O. Irwin (1930), On the frequency distribution of the means of samples from populations of certain of Pearson's types, *Metron*, Vol. 8, No. 4, pp. 51-105. [203]

I. Kaplansky and J. Riordan (1945), Multiple matching and runs by the symbolic

method, *Ann. Math. Stat.*, Vol. 16, pp. 272-277. [154]

T. Kawata and H. Sakamoto (1949), On the characterization of the normal population by the independence of the sample mean and sample variance, *Jour. Math. Soc. Japan*, Vol. 1, pp. 111-115. [208]

J. H. B. Kemperman (1956), Generalized tolerance limits, *Ann. Math. Stat.*, Vol. 27, pp. 180-186. [239]

O. Kempthorne (1952), *The Design and Analysis of Experiments*, John Wiley, New York. [229]

D. G. Kendall (1951), Some problems in the theory of queues, *Jour. Roy. Stat. Soc., Series B*, Vol. 13, pp. 151-185. [192]

D. G. Kendall (1953), Stochastic processes occurring in the theory of queues and their analysis by the method of the Markov chain, *Ann. Math. Stat.*, Vol. 24, pp. 338-354. [192]

D. G. Kendall and K. S. Rao (1950), On the generalized second limit-theorem in the calculus of probabilities, *Biometrika*, Vol. 37, pp. 224-230. [128]

M. G. Kendall (1943 and 1946), *The Advanced Theory of Statistics*, Vol. I (1943), Vol. II (1946) (Vol. I of new three-volume edition appeared in 1958), Charles Griffin, London. [197, 198]

M. G. Kendall (1953), *Rank Correlation Methods*, second edition (first edition, 1948), Charles Griffin, London. [231]

A. Khintchine (1929), Sur la loi des grands nombres, *Comp. Rend. Acad. Sci.*, Vol. 188, pp. 477-479. [111, 250]

A. Kolmogorov (1928), Über die Summen durch den Zufall bestimmer unabhängiger Grössen, *Math. Ann.*, Vol. 99, pp. 309-319. [108]

A. Kolmogorov (1933 *a*), *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergeb. Math. No. 3, Berlin. (English translation by N. Morrison (1950), Chelsea, New York). [1, 10, 26, 99]

邦訳 根本伸司・一条洋共訳 《確率論の基礎概念》 東京図書

R. G. Laha (1954), On a characterization of the gamma distribution, *Ann. Math. Stat.*, Vol. 25, pp. 784-787. [245]

P. S. Laplace (1814), *Théorie Analytique des Probabilités*, second edition, Paris. [145, 202, 253]

P. Lévy (1925), *Calcul des Probabilités*, Gauthier-Villars, Paris. [1, 117, 120]

P. Lévy (1935), Propriétes asymptotiques des sommes des variables aléatoires indépendantes ou enchainees, *Journ. Math. Pures et Appl.*, Vol. 14, p. 347. [253]

P. Lévy (1937), *Théorie de l'addition des variables aléatoires* (third edition, 1954), Gauthier-Villars, Paris. [1, 123]

G. J. Lieberman and D. B. Owen (1961), *Tables of the Hypergeometric Probability Distribution*, Stanford University Press. [136]

J. W. Lindeberg (1922), Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung, *Math. Zeits.*, Vol. 15, pp. 211-225. [252, 253, 254]

M. Loève (1955), *Probability Theory*, D. Van Nostrand, Princeton, N. J. (second edition, 1960, third edition, 1963). [1, 21, 100, 124]

F. M. Lord (1955). Sampling fluctuations resulting from the sampling of test items, *Psychometrika*, Vol. 20, pp. 1-22. [221]

A. J. Lotka (1939), A contribution to the theory of self-renewing aggregates. with special reference to industrial replacement, *Ann. Math. Stat.*, Vol. 10, pp. 1-25. [189]

E. Lukacs (1942), A characterization of the normal distribution, *Ann. Math. Stat.*, Vol. 13, pp. 91-93. [208, 246]

A. Lyapunov (1900), Sur une proposition de la théorie des probabilités, *Bull. de l'Acad. Imp. des Sciences, de St. Petersbourg*, Vol. 13, pp. 359-386. [253]

A. Lyapunov (1901), Nouvelle forme da théoreme sur la limite de probabilité, *Mem. Acad. Science of St. Petersbourg*, Vol. 12, pp. 1-24. [253, 261]

S. Malmquist (1951), On a property of order statistics from a rectangular distribution, *Skand. Aktuar.*, Vol. 33, pp. 214-222. [246]

H. B. Mann (1943), On the construction of sets of orthogonal Latin squares, *Ann. Math. Stat.*, Vol. 14, pp. 401-414. [230]

A. A. Markov (1900), *Wahrscheinlichkeitsrechnung*, Teubner, Leipzig. (German translation of Russian second edition (1908) appeared in 1912; original Russian edition appeared in 1900.) [253]

K. Matusita (1957), Decision rule based on the distance for the classification problem, *Ann. Inst. Stat. Math.*. Vol. 8, pp. 67-77. [248]

P. J. McCarthy (1947), Approximate solutions for means and variances in a certain class of box problems, *Ann. Math. Stat.*, Vol. 18, pp. 349-383. [145]

E. J. McShane and T. Botts (1959), *Real Analysis*, D. Van Nostrand, Princeton, N. J. [21]

R. von Mises (1931), *Wahrscheinlichskeitsrechnung, und ihre Anwendung in der Statistik und Theoretischen Physik*, Deuticke, Leipzig. [10]

E. C. Molina (1942), *Poisson's Exponential Binomial Limit*, D. Van Nostrand, Princeton, N. J. [141]

A. M. Mood (1940), The distribution theory of runs, *Ann. Math. Stat.*, Vol. 11, pp. 367-392. [145, 150]

P. M. Morse (1958), *Queues, Inventories and Maintenance*, John Wiley, New York. [192]

F. C. Mosteller (1046), On some useful "inefficient" statistics, *Ann. Math. Stat.*, Vol. 17, pp. 377-408. [269]

M. E. Munroe (1953), *Introduction to Measure and Integration*, Addison-Wesley, Cambridge, Mass. [1, 27]

National Bureau of Standards (1942), *Tables of Probability Functions*, Vol. 2, New York. [157]

National Bureau of Standards (1949), *Tables of the Binomial Probability Distribution, Appl. Math. Series*, Vol. 6. [139]

J. von Neumann (1950), *Functional Operators*, Vol. 1, Princeton University Press. [50]

J. Neyman (1939), On a new class of "contagious" distributions, applicable in entomology and bacteriology, *Ann. Math. Stat.*, Vol. 10, pp. 35-57. [152]

I. Olkin and J. W. Pratt (1958), A multivariate Chebyshev inequality, *Ann. Math. Stat.*, Vol. 29, pp. 226-234. [113]

E. Parzen (1960), *Modern Probability Theory and its Applications*, John Wiley, New York. [246]

E. S. Pearson and H. O. Hartley (1954), *Biometrika Tables for Statisticians*, Vol. 1, Cambridge University Press. [182, 184]

K. Pearson (1900), On a criterion that a system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling, *Phil. Mag.*, Vol. 50, pp. 157-175. [182, 258]

K. Pearson (1906), On the curves which are most suitable for describing the frequency of random samples of a population, *Biometrika*, Vol. 5, pp. 172-175. [170]

K. Pearson (1920), On the probable errors of frequency constants, Part III, *Biometrika*, Vol. 13, pp. 113-132. [269]

K. Pearson, editor (1922), *Tables of the Incomplete Gamma Function*, Cambridge University Press. [171]

K. Pearson, editor (1934), *Tables of the Incomplete Beta Function*, Cambridge University Press. [173, 186]

E. J. G. Pitman (1937 *a*), The "closest" estimates of statistical parameters, *Proc. Camb. Phil. Soc.*, Vol. 33, pp. 212-222. [245]

S. D. Poisson (1837), *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, precédées des régles générales du calcul des probabilités*, Paris. [140]

A. Renyi (1953), On the theory of order statistics, *Act. Math. Acad. Sci. Hungary*, Vol. 4, pp. 191-231. [246]

P. R. Rider (1955), The distribution of the product of maximum values in samples from a rectangular distribution, *Jour. Amer. Stat. Assn.*, Vol. 50, pp. 1142-1143. [245]

H. Robbins (1944 *a*), On the measure of a random set, *Ann. Math. Stat.*, Vol. 15, pp. 70-74. [246]

H. Robbins (1948), Convergence of distributions, *Ann. Math. Stat.*, Vol. 19, pp. 72-76. [113]

H. Robbins (1954), A remark on the joint distribution of cumulative sums, *Ann. Math. Stat.*, Vol. 25, pp. 614-616. [113]

H. G. Romig (1953), 50-100 *Binomial Tables*, John Wiley, New York. [139]

S. Saks (1937), *Theory of the Integral*, Second Edition (English translation by L. C. Young), Stechert, New York. [21]

I. R. Savage (1962), *Bibliography of nonparametric statistics*, Harvard University Press. Cambridge. [231]

H. Scheffé (1947), A usesul convergence theorem for probability distributions, *Ann. Math. Stat.*, Vol. 18, pp. 434-438. [113]

S. S. Shrikhande (1952), On the dual of some balanced incomplete block designs, *Biometrics*, Vol. 8, pp. 66-72. [229]

N. Smirnov (1935), Über die Verteiling des allgemeinen Gliedes in der Variationsreihe, *Metron*, Vol. 12, No. 2, pp. 59-81. [267, 269]

W. L. Smith (1958), Renewal theory and its ramifications, *Jour. Roy. Stat. Soc., Series B*, Vol. 20, pp. 243-302. [189]

G. W. Snedecor (1937), *Statistical Methods*, Iowa State College Press, Ames, Iowa. [186]

W. L. Stevens (1939), Distributions of groups in a sequence of alternatives, *Ann. Eugen.*, Vol. 9, pp. 10-17. [149]

F. S. Swed and C. Eisenhart (1943), Tables for testing randomness of grouping in a sequence of alternatives, *Ann. Math. Stat.*, Vol. 14, pp. 66-87. [149]

P. C. Tang (1938), The power function of the analysis of variance tests with tables and illustrations of their use, *Stat. Res. Mem.*, University College, London, Vol. 2, pp. 128-149. [244]

T. N. Thiele (1903), *Theory of Observations*, Layton, London (Reprinted in *Ann. Math. Stat.*, Vol. 2, 1931, pp. 165-307). [116]

L. H. C. Tippett (1925), On the extreme individuals and the range of samples take from a normal population, *Biometrika*, Vol. 17, pp. 364-387. [244]

J. W. Tukey (1946), An inequality for deviations from medians, *Ann. Math. Stat.*, Vol. 17, pp. 75-78. [113]

J. W. Tukey (1947), Nonparametric estimation, II. Statistically equivalent blocks and tolerance regions—the continuous case, *Ann. Math. Stat.*, Vol. 18, pp. 529-539. [236]

J. W. Tukey (1948), Nonparametric estimation, III. Statistically equivalent blocks and multivariate tolerance regions—the discontinuous case, *Ann. Math. Stat.*, Vol. 19, pp. 30-39. [239]

J. W. Tukey (1949 *b*), Moments of random group size distributions, *Ann. Math. Stat.*, Vol. 20, pp. 523-539. [153]

J. W. Tukey (1950), Some sampling simplified, *Jour. Amer. Stat. Assn.*, Vol. 45, pp. 501-519. [217, 223]

J. W. Tukey (1956 *a*), Keeping moment-like sampling computations simple, *Ann. Math. Stat.*, Vol. 27, pp. 37-54. [217]

A. Wald (1943), An extension of Wilks' method for setting tolerance limits, *Ann. Math. Stat.*, Vol. 14, pp. 45-55. [236]

A. Wald and J. Wolfowitz (1944), Statistical tests based on permutations of the observations, *Ann. Math. Stat.*, Vol. 15, pp. 358-372. [261, 262]

D. L. Wallace (1958), Asymptotic approximations to distributions, *Ann. Math. Stat.*, Vol. 29, pp. 635-654. [261]

J. E. Walsh (1946), Some order statistic distributions for samples of size 4, *Ann. Math. Stat.*, Vol. 17, pp. 246-248. [249]

E. T. Whittaker and G. N. Watson (1927), *A Course in Modern Analysis* (fourth edition), Cambridge University Press. [119, 176]

D. V. Widder (1947), *Advanced Calculus*, Prentice-Hall, New York. [58]

S. S. Wilks (1959 *b*), Recurrence of extreme observations, *Jour. Austral. Math. Soc.*, Vol. 1, pp. 106-112. [247]

J. D. Williams (1946), An approximation to the probability integral, *Ann. Math. Stat.*, Vol. 17, pp. 363-365. [187]

G. U. Yule (1924), A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis, *Phil. Trans. Roy. Soc., London, Series B*, Vol. 213, pp. 21-87. [191]

索　引

タ 行

ナ 行

ハ 行

マ 行

訳 者 紹 介

田中英之（たなか ひでゆき）　1962年　東京教育大学理学部数学科卒業
現　在　味の素株式会社勤務

岩本誠一（いわもと せいいち）　1970年　九州大学大学院理学研究科修士課程修了
現　在　福岡大学理学部講師

数 理 統 計 学 1　¥ 2000

1971 年 12 月 28 日　第 1 刷発行

検印省略

著 者　S. S. ウィルクス

訳 者　田 中 英 之
岩 本 誠 一

発行所　東 京 図 書 株 式 会 社
本 社　東京都新宿区本塩町 23（タナカビル）
振替 東京 13803 電話 (357) 0636～7
営業所　東 京 都 文 京 区 水 道 1－8－1
電話 (814) 7818～9

落丁乱丁はお取替えします

3341-2056-5160